

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

Searcher: _____

Searcher Phone #: **if Contact: 308-4499**
Sheppard

Searcher Location: _____

Date Searcher Picked Up: **8/8/01**

Date Completed: _____

Searcher Prep & Review Time: _____

Clerical Prep Time: _____

Online Time: _____

Type of Search

NA Sequence (#) _____ STN _____

AA Sequence (#) _____ Dialog _____

Structure (#) _____ Questel/Orbit _____

Bibliographic _____ Dr.Link _____

Litigation _____ Lexis/Nexis _____

Fulltext _____ Sequence Systems _____

Patent Family _____ WWW/Internet _____

Other _____ Other (specify) _____

Vendors and cost where applicable

L3 ANSWER 1 OF 5 CAPLUS COPYRIGHT 2003 ACS
 AN 2003:117556 CAPLUS
 TI Method for treating diabetes and obesity with fibroblast growth factor 21
 IN Glasebrook, Andrew Lawrence; Hammond, Lisa Janine; Kharitononkov, Alexei;
 Shiyanova, Tatiyana Leonidovna
 PA Eli Lilly and Company, USA
 SO PCT Int. Appl., 32 pp.
 CODEN: PIXXD2

DT Patent
 LA English

FAN.CNT 1

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	WO 2003011213	A2	20030213	WO 2002-US21290	20020722
	W:	AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
	RW:	GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG			
PRAI	US 2001-308702P	P	20010730		
	US 2002-347991P	P	20020110		
AB	A method for treating diabetes and obesity comprises administering an effective amt. of fibroblast growth factor 21. FGF-21 stimulated glucose uptake in a mouse obesity model and reduced plasma glucose levels in a dose-dependent manner in fat, diabetic ZDF rats. FGF-21 did not induce hypoglycemia in lean ZDF rats.				

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 6, 2001, 20:58:04 ; Search time 1964.36 Seconds
(without alignments)
5726.494 Million cell updates/sec

Title: US-09-391-861-1
Perfect score: 1190
Sequence: 1 gaggatccagccgaagagg.....aaaaaaaaaaaaaaaaaaaa 1190

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues 20456230
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1:	gb_est1:*
2:	gb_est2:*
3:	gb_est3:*
4:	gb_est4:*
5:	gb_est5:*
6:	gb_est6:*
7:	gb_est7:*
8:	gb_est8:*
9:	gb_est9:*
10:	gb_est10:*
11:	gb_est11:*
12:	gb_est12:*
13:	gb_est13:*
14:	gb_est14:*
15:	gb_est15:*
16:	gb_est16:*
17:	gb_est17:*
18:	gb_est18:*
19:	gb_est19:*
20:	gb_est20:*
21:	gb_est21:*
22:	gb_est22:*
23:	gb_est23:*
24:	gb_est24:*
25:	gb_est25:*
26:	gb_est26:*
27:	gb_est27:*
28:	gb_est28:*
29:	gb_est29:*
30:	gb_est30:*
31:	gb_est31:*
32:	gb_est32:*
33:	gb_est33:*
34:	gb_est34:*
35:	gb_est35:*
36:	gb_est36:*
37:	gb_est37:*
38:	gb_est38:*
39:	gb_est39:*
40:	gb_est40:*
41:	gb_est41:*
42:	gb_est42:*
43:	gb_est43:*
44:	gb_est44:*
45:	gb_est45:*
46:	gb_est46:*
47:	gb_est47:*
48:	gb_est48:*
49:	gb_est49:*
50:	gb_est50:*
51:	gb_est51:*
52:	gb_est52:*
53:	gb_est53:*
54:	gb_est54:*
55:	gb_est55:*
56:	gb_est56:*
57:	gb_est57:*
58:	gb_est58:*
59:	gb_est59:*
60:	gb_est60:*
61:	gb_est61:*
62:	gb_est62:*
63:	gb_est63:*
64:	gb_est64:*
65:	gb_est65:*
66:	gb_est66:*
67:	gb_est67:*
68:	gb_est68:*
69:	gb_est69:*
70:	gb_est70:*
71:	gb_est71:*
72:	gb_est72:*
73:	gb_est73:*
74:	gb_est74:*
75:	gb_est75:*
76:	gb_est76:*
77:	gb_est77:*
78:	gb_est78:*
79:	gb_est79:*
80:	gb_est80:*
81:	gb_est81:*
82:	gb_est82:*
83:	gb_est83:*
84:	gb_est84:*
85:	gb_est85:*
86:	gb_est86:*
87:	gb_est87:*
88:	gb_est88:*
89:	gb_est89:*
90:	gb_est90:*
91:	gb_est91:*
92:	gb_est92:*
93:	gb_est93:*
94:	gb_est94:*
95:	gb_est95:*
96:	gb_est96:*
97:	gb_est97:*
98:	gb_est98:*
99:	gb_est99:*
100:	gb_est100:*
101:	gb_est101:*
102:	gb_est102:*
103:	gb_est103:*
104:	gb_est104:*
105:	gb_est105:*
106:	gb_est106:*
107:	gb_est107:*
108:	gb_est108:*
109:	gb_est109:*
110:	gb_est110:*
111:	gb_est111:*
112:	gb_est112:*
113:	gb_est113:*
114:	gb_est114:*
115:	gb_est115:*
116:	gb_est116:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.


```

SOURCE
ORGANISM      human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: DTD/DTP/Gazdar
              cDNA Library Preparation: Life Technologies, Inc.
              DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
              Clone distribution: Incyte Genomics, Inc.
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: L1AM9649 row: e column: 06
              High quality sequence stop: 694.
FEATURES
source
Location/Qualifiers
1..775
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3881069"
/clone_lib="NIH_MGC_68"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.8 Kb. Library constructed by Life
Technologies."
BASE COUNT      182 a 219 c 205 g 169 t
ORIGIN
Query Match      18.5%; Score 220.2; DB 140; Length 775;
Best Local Similarity 87.3%; Pred. No. 1.2e-28;
Matches 310; Conservative 0; Mismatches 38; Indels 7; Gaps 6;
QY 1 gagatccagcgaagagagcagcactcagccactcagccactcagctcactcactcagcaga 60
DB 422 GAGGATCCAGCCGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 481
QY 61 actggaatctggcaccattcgaacactcagctcctcagcagctcacaccccgagat- 119
DB 482 ACTGGAATCTGGCACCACCAATTTCTAAA-CACTCAGCTTCTCGAGCTCACACCCCGGAGATC 540
QY 120 caccctgagcaccgagccattgatgactcggacgagacccgggttcgagcactcaggact 179
DB 541 CACCTGAGGACCCGAGCCATTGATGGACTCGGACGAGACCGGGTTGAGCACTTCAGGGCT 600
QY 180 gt-gggtttctgtgctgtgtcttctgtcgtgagcctcgcagggcacacccca-tccct 237
DB 601 GTGGGTATCTGTGCTGTGTGTGTCTGTGCTGGAGCCCTGCCAGTGAACCCCATTTCCCT 660
QY 238 gactccagtcct-ctctgcgaattcggggcccaagtcccgagcgggtacactctacacaga 296
DB 661 GAATCCAGTCTCTCTCTCTGCAATTCGGGGGCAAGTCCCGGAGGAGGAGGAGGAGGAGGAG 720
QY 297 tgaatgccacgacagagaagccacactggagatcagggaggatggagacggtgggg 351
DB 721 TAATGCCCA--AAAAAAGAGACCACCTGGAGATCAAGGCGGATTCGACACAGGGGG 773

RESULT 4
AV049138
LOCUS      AV049138 307 bp mRNA EST 22-JUN-1999
DEFINITION AV049138 Mus musculus pancreas C57BL/6J adult Mus musculus cDNA
clone 1810008N24, mRNA sequence.
ACCESSION AV049138
VERSION    AV049138.1 GI:5134910
KEYWORDS   EST.
SOURCE     house mouse.

```

```

ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS      Carninci, P., Shiyama, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,
              Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara
              A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M., Kawai, J.,
              Kikuchi, N., Kojima, Y., Matsuyama, T., Niitsuma, H., Oda, H., Owa, C.,
              Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, F., Sogabe, Y., Sugahara
              Y., Suzuki, H., Suzuki, H., Tateno, M., Tomaru, Y., Tominaga, N.,
              Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M.,
              Yurimoto, M., Okazaki, Y. and Hayashizaki, Y.
              RIKEN Mouse ESTs
              Unpublished (1999)
              Contact: Chie Owa
              Genome Science Laboratory
              RIKEN
              3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
              Tel: 81-298-36-9145
              Fax: 81-298-36-9098
              Email: genome-res@rtc.riken.go.jp
              Thermostabilization and thermoactivation of thermolabile enzymes by
              trehalose and its application for the synthesis of full length cDNA
              (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
              Transcriptional sequencing: A method for DNA sequencing using RNA
              polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
              Please visit our web site (http://genome.rtc.riken.go.jp) for
              further details.
FEATURES
source
Location/Qualifiers
1..307
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="1810008N24"
/clone_lib="Mus musculus pancreas C57BL/6J adult"
/sex="male"
/tissue_type="pancreas"
/dev_stage="adult"
BASE COUNT      65 a 72 c 67 g 103 t
ORIGIN
Query Match      10.7%; Score 127; DB 109; Length 307;
Best Local Similarity 68.9%; Pred. No. 1.6e-12;
Matches 210; Conservative 0; Mismatches 80; Indels 15; Gaps 2;
QY 578 tgcagggaacaagtccccaacccgggacccctgcaccccgagcagcagcagcagcagcagc 637
DB 1 TGCCTTAGAAGGAATCCCAACACAGGATGCAACATTCCTGGGAGCTGTGGCTTCCTGC 60
QY 638 cactaccaggcctccccccaccccgagcagcagcagcagcagcagcagcagcagcagcagc 697
DB 61 CCATGGCAGGCTGATTCATGAGCCCCCAAGACCAAGCAGGATTCCTGCCCCAGAGCCCC 120
QY 698 ccatgtgggtcctcggaccctcagcctcagcctcagcctcagcctcagcctcagcctcagcctc 757
DB 121 CAGATGTGGATCTTTTGACCCCTTGAGCTGTAGAGCTTACAGGCGCGAAGCCCCA 180
QY 758 gtaagcttcctga-----agccagaggtgtttactatgacatcctcttatt 808
DB 181 GCTATGCGCTCTGATTTTCTCTGAATTTAGGCTGTTTTTTTTTTGGGTTTCCACTATT 240
QY 809 tattaggtattattctattattt-----ttttattttcttacttgagataataa 862
DB 241 ATTACGGGTATTTATTTTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 300
QY 863 gaggtt 867
DB 301 GAGTT 305
RESULT 5
AV050323

```



```

|||||
Db 187 TGTTCCTTTTGGGTTTCCACTTATTATTACGGGTATTATTATTATTATTATTATTAG 246
QY 836 ttatctttcttactgagataaagagcttccagagaggat 879
|||||
Db 247 TTTTCTTTCTTACTTGGAAATAAAGAGCTCTGAAGAAAAAT 290
|||||

RESULT 7
AV052213 288 bp mRNA EST 22-JUN-1999
LOCUS AV052213 Mus musculus pancreas C57BL/6J adult Mus musculus cDNA
DEFINITION clone 1810021E04, mRNA sequence.
ACCESSION AV052213
VERSION AV052213.1 GI:5137985
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 288)
Carninci,P., Shibata,K., Ozawa,Y., Konno,H., Itoh,M., Aizawa,K.,
Akahira,S., Akiyama,J., Fukuda,S., Fukunishi,Y., Funayama,T., Hara
A., Hayatsu,N., Hori,F., Ishikawa,T., Itoh,M., Izawa,M., Kawai,J.,
Kikuchi,N., Kojima,Y., Matsuyama,T., Naitsuma,H., Oda,H., Owa,C.,
Sato,K., Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara
Y., Suzuki,H., Suzuki,H., Tateo,M., Tomaru,Y., Tomimaga,N.,
Watanabe,S., Yagame,M., Yamamura,T., Yokota,T., Yoshino,M.,
Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
RIKEN Mouse ESTs
Unpublished (1999)
Contact: Chie Owa
Genome Science Laboratory
RIKEN
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145
Fax: 81-298-36-9098
Email: genome-res@rtc.riken.go.jp
Thermostabilization and thermostabilization of thermolabile enzymes by
trehalose and its application for the synthesis of full length cDNA
(Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (http://genome.rtc.riken.go.jp) for
further details.
FEATURES
source
Location/Qualifiers
1..288
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="1810021E04"
/clone_lib="Mus musculus pancreas C57BL/6J adult"
/sex="male"
/tissue_type="pancreas"
/dev_stage="adult"
BASE COUNT 62 a 65 c 61 g 100 t
ORIGIN
Query Match 9.9%; Score 118; DB 109; Length 288;
Best Local Similarity 70.6%; Pred. No. 5.8e-11;
Matches 192; Conservative 0; Mismatches 65; Indels 15; Gaps 2;

QY 623 cagctgcttctgcactaccaggctgccccccgcagcccgagccaccggaatcc 682
|||||
Db 1 CTGTGGCTTCTGCGCATGCGCGCTGATCCAGAGCCGCCACAGCAGCAGGATGCC 60
|||||
QY 683 tggccccagccccccagtgctgggctctcgaccccttgagcatggtggagctccc 742
|||||
Db 61 TGCCTCCAGAGCCCGCAGATGTGGTTCCTGTGACCCCTGAGCTGTAGAGCCTTAC 120
|||||
QY 743 agggccgagcccgactacgttctctga-----agccagaggctgttactatga 793
|||||

```

```

Db 121 AGGCCGGAAGCCCGAGCTATGGCTCTGATTTTCTGAAATTAGGCTGTTTTTTTGG 180
QY 794 catctctcttattattatttaggtatttatttatttattt-----ttttattttttt 847
|||||
Db 181 GGTTCCTCCACTTATTATTACGGGTATTATTATTATTATTATTATTATTATTATTCTT 240
|||||
QY 848 acttgagataataagaggttccagagaggat 879
|||||
Db 241 ACTTGGATATATAAGAGCTTTTGAAGAAAAAT 272
|||||

RESULT 8
BG389761 866 bp mRNA EST 12-MAR-2001
LOCUS BG389761 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4523457 5',
DEFINITION mRNA sequence.
ACCESSION BG389761.1 GI:13283197
VERSION BG389761
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 866)
NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10426 row: C column: 10
High quality sequence stop: 573.
FEATURES
source
Location/Qualifiers
1..866
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4523457"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-qt primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 350 a 141 c 223 g 152 t
ORIGIN
Query Match 9.7%; Score 115.6; DB 153; Length 866;
Best Local Similarity 89.9%; Pred. No. 1.4e-10;
Matches 124; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1053 tgcactcagcccgagccagagcgagattccatctcaaaaaataataataaa 1112
|||||
Db 417 TGCACCTCAGCTGGCAACAGAGAGCTCCATCTCAAAAAAATAAAAAA 476
|||||
QY 1113 ataataataataataaaaaaataaaaaaataaaaaaataaaaaa 1172
|||||
Db 477 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 536
|||||
QY 1173 aaaaaaataaaaaaataaaaaaataaaaaaataaaaaa 1190
|||||
Db 537 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 554
|||||

RESULT 9

```


[illegible]

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AM9508 row: n column: 09

High quality sequence stop: 223.

Location/Qualifiers

FEATURES

source

1. .1029

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4186952"

/lab_host="NCI_CGAP_Brn64"

/tissue_type="glioblastoma with EGFR amplification"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: brain; Vector: PCMV-SPORT6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.57 kb. Constructed by Life

Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 432 a 241 c 223 g 133 t

ORIGIN

Query Match

Best Local Similarity 9.2%; Score 109; DB 147; Length 1029;

Matches 121; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1050 ctgtgcactccagccagcgccagcgagattccatctcaaaaaataataataaa 1109

1109 | ||||| ||||| || ||||| ||||| || ||||| ||||| || ||||| || |||||

Db 53 CACTGCACCTCCGCTGGGCACAGAGTGAGACTCTGTCTCAAAAAAAAAAAAAAT 112

QY 1110 taaataataataataataataataataataataataataataataataaa 1169

1169 | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 113 AA 172

QY 1170 aaaaaaaaaaaaaaaaaaaaaa 1190

1190 | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 173 AAAAAAAAAAAAAAAAAAAAAA 193

Search completed: August 6, 2001, 21:43:02

Job time: 2698 sec

Result No.	Score	Query Match	Length	DB	ID	Description	
1	92.2	7.7	2082	2	US-08-785-310A-2	Sequence 2, Appli	
2	88.2	7.4	1882	4	US-09-370-253-1	Sequence 1, Appli	
3	88	7.4	1066	1	US-08-157-101A-4	Sequence 4, Appli	
4	88	7.4	1733	3	US-09-073-569-1	Sequence 1, Appli	
5	88	7.4	6671	1	US-08-240-443-1	Sequence 1, Appli	
6	88	7.4	6671	1	US-08-457-459-1	Sequence 1, Appli	
7	88	7.4	6671	1	US-08-555-678-1	Sequence 1, Appli	
8	88	7.4	6671	5	PCT-US95-02275-1	Sequence 1, Appli	
9	87.2	7.3	140	1	US-08-628-417-5	Sequence 5, Appli	
10	87.2	7.3	240	1	US-08-628-417-6	Sequence 6, Appli	
11	87.2	7.3	1117	4	US-09-247-373B-33	Sequence 33, Appli	
12	87	7.3	1474	4	US-08-821-994-64	Sequence 64, Appli	
13	86.4	7.3	144	1	US-08-702-344-26	Sequence 26, Appli	
14	85.8	7.2	2447	2	US-09-014-969-14	Sequence 14, Appli	
15	85.6	7.2	111	4	US-09-297-535-23	Sequence 23, Appli	
16	85.6	7.2	117	1	US-08-702-344-3	Sequence 3, Appli	
17	85.6	7.2	121	4	US-09-297-535-20	Sequence 20, Appli	
18	85.6	7.2	9589	1	US-07-925-695-1	Sequence 1, Appli	c
19	85.6	7.2	9589	1	US-07-925-695-2	Sequence 2, Appli	c
20	83.6	7.0	98	1	US-08-088-658-42	Sequence 42, Appli	
21	83.6	7.0	98	2	US-08-471-907A-42	Sequence 42, Appli	
22	82.2	6.9	903	5	PCT-US95-06406A-21	Sequence 21, Appli	
23	80.8	6.8	260	1	US-08-520-678A-29	Sequence 29, Appli	c
24	80.6	6.8	1641	1	US-08-300-903A-8	Sequence 8, Appli	
25	80.4	6.8	6243	2	US-09-056-075-1	Sequence 1, Appli	
26	80.2	6.7	333	4	US-09-018-584A-27	Sequence 27, Appli	
27	80.2	6.7	1582	3	US-08-545-196B-10	Sequence 10, Appli	

FILING DATE: 05-APR-1994
REGISTRATION NUMBER: 03/0

```

; STRANDEDNESS: single

```

Query Match 7.4%; Score 88; DB 1; Length 6671;
Best Local Similarity 90.4%; Pred. No. 1.8e-08;
Matches 94; Conservative 0; Mismatches 10; Indels

Dy
Db

1086 atctcaaaaaataaataaaataaaaataataataaaaaaaaaa 1145
||| ||||| ||| ||| ||| ||| ||| ||| |||||
6568 ATCTGAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAA 6627

Dy	1146	aaaaaaaaaaaaaaaaaaaaaaaaaaaa	1189
Db	6628	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	6671

```

RESULT      7
US-08-555-678-1
; Sequence 1, Application US/08555678
; Patent No. 5763174
; GENERAL INFORMATION:
; APPLICANT: Nishikura, Kazuko
; TITLE OF INVENTION: RNA Editing Enzyme and Methods
; TITLE OF INVENTION: of Use Thereof
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/555,678
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,794
; FILING DATE: 17-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/280,443
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/457,459
; FILING DATE: 01-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST49DUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9206
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6671 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 155..3832
US-08-555-678-1

Query Match          7.4%; Score 88; DB 1; Length 6671;
Best Local Similarity 90.4%; Pred. No. 1.8e-08;
Matches    94; Conservative   0; Mismatches 10; Indels   0; Gaps   0;

Qy  1086 atctcaaaaaataataatataaatataataataataataataataataataaa 1145
     |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  6568 ATCTGAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 6627
     |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

Qy  1146 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1189
     |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  6628 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 6671
     |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT      8
PCT-US95-02275-1
; Sequence 1, Application PCT/TUS9502275
; GENERAL INFORMATION:
; - APPLICANT: Wistar Institute of Anatomy & Biology
; TITLE OF INVENTION: RNA Editing Enzyme and Methods of Use
```


FILING DATE: 435
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: BIFFONI, ULYSSES J
 REGISTRATION NUMBER: 39,908
 REFERENCE/DOCKET NUMBER: DAM 398-94
 TELECOMMUNICATION INFORMATION:

1067	ggccacagcgcgagattccatctccaaaaaaataataataataataataataata	1156
QY		
Db		
971	ggcaataatcataattctttaaataaaaaataaaaaataaaaaataaaaa	1030
QY		
1127	aaataaaaaataaaaaataaaaaataaaaaataaaaaataaaaaataaaaa	1186
QY		
1031	aaaaataaaaaataaaaaataaaaaataaaaaataaaaaataaaaaataaaaa	1090
Db		
QY		
1187	aaaa 1190	
QY		

```
Db 1091 aaca 1094
RESULT 12
US-08-821-994-64
; Sequence 64, Application US/08821994A
; Patent No. 6228643
; GENERAL INFORMATION:
; APPLICANT: Greenland, Andrew J
; APPLICANT: Thomas, Didier RP
; APPLICANT: Jepson, Ian
; TITLE OF INVENTION: Promoters
; FILE REFERENCE: PPD 50108
; CURRENT APPLICATION NUMBER: US/08/821,994A
; CURRENT FILING DATE: 1997-03-22
; EARLIER APPLICATION NUMBER: PCT/GB97/00729
; EARLIER FILING DATE: 1997-03-18
; EARLIER APPLICATION NUMBER: GB 9606062.9
; EARLIER FILING DATE: 1996-03-22
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Brassica napus
US-08-821-994-64

Query Match 7.3%; Score 87; DB 4; Length 1474;
Best Local Similarity 90.3%; Pred. No. 2e-08;
Matches 93; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1087 tctcaaaaaataaataaataaataaataaataaataaataaataaataaataaataa 1146
Db 1372 tctcaaaaaataaataaataaataaataaataaataaataaataaataaataa 1431

QY 1147 aaaaaataaataaataaataaataaataaataaataaataaataaataa 1189
Db 1432 aaaaaataaataaataaataaataaataaataaataaataaataaataa 1474

RESULT 13
US-08-702-344-26
; Sequence 26, Application US/08702344
; Patent No. 5723315
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: LaVallie, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/08/702,344
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sprunger, Suzanne A.
; REGISTRATION NUMBER: 41,323
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8284
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA

Db 1091 aaca 1094
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 144 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-702-344-26

Query Match 7.3%; Score 86.4; DB 1; Length 144;
Best Local Similarity 89.4%; Pred. No. 1.5e-08;
Matches 93; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1087 tctcaaaaaataaataaataaataaataaataaataaataaataaataa 1146
Db 19 TTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 78

QY 1147 aaaaaataaataaataaataaataaataaataaataaataaataaataa 1190
Db 79 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 122

RESULT 14
US-09-014-969-14
; Sequence 14, Application US/09014969
; Patent No. 5965397
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Racie, Lisa A.
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; APPLICANT: Agostino, Michael J.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/09/014,969
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sprunger, Suzanne A.
; REGISTRATION NUMBER: 41,323
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8284
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 6, 2001, 21:09:14 ; Search time 163 Seconds
(without alignments)
4584.067 Million cell updates/sec

Title: US-09-391-861-1

Perfect score: 1190

Sequence: 1 gagatccagccgaagagg.....aaaaaaaaaaaaaaaaaaaa 1190

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_0601.*
1: /SID88/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	495	41.6	514	22	AAF76715
2	401.8	33.8	659	21	AAAF75631
3	107.4	9.0	1457	21	AAAG1287
4	95.2	8.0	1376	20	AAAX04325
5	94.8	8.0	875	21	AAAC59201
6	94.6	7.9	3213	20	AAV55742
7	94	7.9	667	21	AAAC59098
8	92.8	7.8	1722	21	AACT71124
9	92.4	7.8	1529	21	AAAC59838
10	92.2	7.7	2082	19	AAV41257
11	91.8	7.7	1985	21	AAAC59406

12	91.6	7.7	936	20	AAZ06226	Human secreted pro
13	91.6	7.7	2753	22	AAAF91862	Human secreted pro
14	91.4	7.7	396	22	AAAF94862	Human ovarian can
15	91	7.6	1259	13	AAQ23028	Clone W264 encodin
16	91	7.6	1388	21	AAQ59295	Human secreted pro
17	90.8	7.6	550	19	AAV41916	Nucleotide sequenc
18	90.8	7.6	905	21	AAAG4642	Partial sequence M
19	90.4	7.6	1733	21	AAA37036	Human PRO1411 (UNQ
20	90.4	7.6	1733	22	AAAF54238	DNA encoding prote
21	90.4	7.6	1734	22	AAAF92083	Human PRO1411 cDNA
22	90.4	7.6	1734	22	AAAC87038	Nucleotide sequenc
23	89.8	7.5	3508	21	AAA16619	Human secreted pro
24	89.8	7.5	4909	22	AAAF24165	Human secreted pro
25	89.6	7.5	882	21	AAZ52528	Human secreted pro
26	89.6	7.5	909	21	AACT7961	Human secreted pro
27	89.6	7.5	1690	21	AAZ90632	Human adipose tiss
28	89.6	7.5	2186	22	AAAF73412	Grand fir monoterp
29	89.4	7.5	876	21	AAAC63439	Human fetal brain
30	89.4	7.5	1936	21	AAZ43798	Human gene express
31	89.2	7.5	300	20	AAZ14372	Coding sequence fo
32	89.2	7.5	772	19	AAV58363	Partial sequence M
33	89.2	7.5	801	21	AAAG4638	Human ZALPHA29 DNA
34	89.2	7.5	813	22	AAAF31470	Human secreted pro
35	89.2	7.5	959	21	AAZ26370	Human secreted pro
36	89.2	7.5	1149	21	AAAC81030	Human secreted pro
37	89.2	7.5	2496	19	AAV63189	CDNA from clone cr
38	89.2	7.5	7753	15	AAQ56826	Norwalk virus stra
39	89	7.5	396	22	AAAF94842	Human ovarian can
40	89	7.5	441	20	AAZ09474	p135-NT3 construct
41	89	7.5	872	22	AAAF31060	Rat clone 70129147
42	89	7.5	1495	21	AAZ27985	Corn CCR4 transcri
43	89	7.5	1751	21	AAAC68122	Human secreted pro
44	89	7.5	4120	20	AAZ09473	Human RNA helicase
45	88.8	7.5	2846	21	AAZ65034	Membrane-bound pro

ALIGNMENTS

RESULT 1

AAF76715
ID AAF76715 standard; DNA; 514 BP.

XX AAF76715;

XX 17-MAY-2001 (first entry)

DE Human fibroblast growth factor 19 homologue FGF19X coding sequence.
XX Human; fibroblast growth factor 19; FGF19X; embryonic development;
KW proliferative disorder; cancer; restenosis; psoriasis;
KW rheumatoid arthritis; Dupuytren's contracture; ds.

XX Homo sapiens.

XX WO200118209-A1.

XX 15-MAR-2001.

PF 08-SEP-2000; 2000WO-US24863.

PR 10-SEP-1999; 99US-0153303.

PR 03-APR-2000; 2000US-0194246.

PR 19-JUL-2000; 2000US-0619251.

PA (CURA-) CURAGEN CORP.

PI Shimkets RA, Vernet C, Burgess C, Fernandes E, Taupier R;

PI Quinn KE, Spyteck KA, Rastelli L, Herrmann JL;

DR WPI; 2001-218559/22.

DR P-PSDB; AAB/3069.

XX

PT Fibroblast growth factor-19X polypeptides and polynucleotides useful
PT for diagnosis, prevention, treatment of proliferative, differentiative,
PT tumorigenic disorders such as tumor, psoriasis, rheumatoid arthritis -

XX Claim 3; Page 7; 97pp; English.

XX The present invention describes the protein and coding sequences of the
CC human fibroblast growth factor (FGF) 19 homologue FGF19X, the latter of
CC which is shown here. The sequences can be used in the prognosis and
CC treatment of proliferative diseases such as cancer, restenosis,
CC psoriasis, rheumatoid arthritis and Dupuytren's contracture, as well as
CC to stimulate cell growth for treating neurological disorders such as
CC Alzheimer's disease.

XX Sequence 514 BP; 103 A; 157 C; 155 G; 99 T; 0 other;

Query Match 41.6%; Score 495; DB 22; Length 514;
Best Local Similarity 100.0%; Pred. No. 3e-64;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 cagcttctccgagctcacaccccgagatcacctgagaccgagccattgatgactcg 150
DB 1 cagcttctccgagctcacaccccgagatcacctgagaccgagccattgatgactcg 60
QY 151 gacgagacgggttcgagcactcagactgtgggtttctgtgctggtcttctgctg 210
DB 61 gacgagacgggttcgagcactcagactgtgggtttctgtgctggtcttctgctg 120
QY 211 ggaagctgacgagcagacccctcctgactccactcctcctcctcctcctcctcctc 270
DB 121 ggaagctgacgagcagacccctcctgactccactcctcctcctcctcctcctcctc 180
QY 271 gtcggcagcgttacctctacacagatgacccgagcagacagccacactggagatc 330
DB 181 gtcggcagcgttacctctacacagatgacccgagcagacagccacactggagatc 240
QY 331 aaggagatgagcagcgttggggcgctgctgacagagcccgaaagtctcctcagctg 390
DB 241 aaggagatgagcagcgttggggcgctgctgacagagcccgaaagtctcctcagctg 300
QY 391 aagccttgaagcgggagttatcaaatcttgggagtcacagacatccaggctcctgtgc 450
DB 301 aagccttgaagcgggagttatcaaatcttgggagtcacagacatccaggctcctgtgc 360
QY 451 cagcggcagatggggcctgtatgagctcacttgcactttgacccctgagcctgcagcttc 510
DB 361 cagcggcagatggggcctgtatgagctcacttgcactttgacccctgagcctgcagcttc 420
QY 511 cgggagctgtcttgagcagcagacatacaatgtttaccagtcgagccacggtccctccg 570
DB 421 cgggagctgtcttgagcagcagacatacaatgtttaccagtcgagccacggtccctccg 480
QY 571 ctgacactgccaggg 585
DB 481 ctgacactgccaggg 495

RESULT 2

AAA75631
ID AAA75631 standard; DNA; 659 BP.
XX AC AAA75631;
XX XX
DT 22-JAN-2001 (first entry)
XX DE
XX Nucleotide sequence of murine fibroblast growth factor (FGF)-21.
KW pd10-VEGFuc; gene delivery vector; eye disease; neovascular disease;
KW neurotrophic factor; anti-angiogenic factor; eye disease; glaucoma;
KW macular degeneration; diabetic retinopathy; retinitis pigmentosa;
KW inherited retinal degeneration; surgery-induced retinopathy;
KW retinal detachment; photic retinopathy; toxic retinopathy;

trauma-induced retinopathy; wet age related macular degeneration;
ARMD; retinopathy; fibroblast growth factor-20; FGF-20; ss.

Mus sp.

Key Location/Qualifiers
CDS 14..646
FT /*tag= a
FT /product= "Fibroblast growth factor 21"

W0200054813-A2.

21-SEP-2000.

15-MAR-2000; 2000WO-US07062.

15-MAR-1999; 99US-0124460.

06-JAN-2000; 2000US-0174984.

(CHIR) CHIRON CORP.

(REGC) UNIV CALIFORNIA.

Manning WC, Dwarki VJ, Rendahl K, Zhou S, McGee LH, Lau D;

Flannery JG, Miller S, Wang F, Di Polo A;

WPI; 2000-618862/59.

P-PSDB; AAB18636.

Treating or preventing eye diseases or inhibiting neovascular disease
of the eye, comprises intraocularly administering a gene delivery
vector that directs expression of neurotrophic factors or
anti-angiogenic factors -

Disclosure; Fig 30; 86pp; English.

The present sequence encodes a murine fibroblast growth factor (FGF)
-21. FGF-20 is expressed using a gene delivery vector of the invention.
Vectors of the invention are used for treating or preventing eye
diseases, or inhibiting neovascular disease of the eye. The gene
delivery vector directs the expression of one or more neurotrophic
factors, or anti-angiogenic factors, such that the disease of the eye
is prevented or treated. The gene delivery vectors are useful for
treating or preventing diseases of the eye such as macular degeneration,
diabetic retinopathy, inherited retinal degeneration such as retinitis
pigmentosa, glaucoma, surgery-induced retinopathy, retinal detachment,
photic retinopathy, toxic retinopathy or trauma-induced retinopathy and
for inhibiting neovascular diseases such as wet age related macular
degeneration (ARMD) or retinopathy of prematurity.

Sequence 659 BP; 140 A; 208 C; 179 G; 132 T; 0 other;

Query Match 33.8%; Score 401.8; DB 21; Length 659;
Best Local Similarity 80.7%; Pred. No. 1.1e-50;
Matches 469; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 191 tgcctgctggtctctctgctggagcctccagccacaccccatccctgactccagctc 250
DB 66 tgcctgctgctctctgctgggtctaccaagcaccatccctgactccagccccc 125
QY 251 tcctgcaattcgggggccaaagtcgagcgggtacctctacacagatgagccacagaga 310
DB 126 tcctccagtttgggggtcaagtcgagcgggtacctctacacagatgagccacagaca 185
QY 311 cagaagccacactgagatcaggagatgggacggtggggcgctgctgacacagagcc 370
DB 186 ctgaagccacactgagatcaggagatgggagagtggaacagtggtgagcagacacgcagtc 245
QY 371 ccgaagctctctgctcagctgaaagccttgaaagccgggaggtattcaaatcttgggagtc 430
DB 246 cagaagctctctgctcagctgaaagccttgaaagccggggtcattcaaatcttgggagtc 305
QY 431 agacatccaggttctctgtgcccagcggccagatggggccctgtatggtcgtccacttg 490

CC		and include products for the diagnosis or treatment of cancer, tumours,					
CC	AIDS, autoimmune disorders, allergy, cardiovascular disorders, viral,						
CC	bacterial and fungal infection. The genes are used to generate fusion						
CC	proteins by linking to the gene a human immunoglobulin portion (AAA61251)						
CC	for increasing stability of the fused protein as compared to the						
CC	secreted protein only.						
XX							
SQ	Sequence 1457 BP; 440 A; 340 C; 307 G; 370 T; 0 other;						
	Query Match	9.0%;	Score 107.4;	DB 21;	Length 1457;		
	Best Local Similarity	80.3%;	Pred. No. 7.6e-08;				
	Matches 126; Conservative	0;	Mismatches 31;	Indels	0;	Gaps	
Qy	1034 ttgagcggagtctgcgtctgcactccacgccaggccacagcgagatccatctcaaa	1093					
Db							
Db	1276 tttagtcttgctgacctgcacgcgcctgggcaacaagagcgaaactctgtcaaa	1335					
Qy	1094 aaataaataataataataataataataataataataataataataataaaaaaaa	1153					
Db							
Db	1336 aaaaaaiaaa	1395					
Qy	1154 aa	1190					
Db							
Db	1396 aa	1432					
RESULT	4						
AAX04325							
ID	AAX04325 standard; DNA; 1376 BP.						
XX							
AC	AAX04325;						
XX							
DT	13-APR-1999 (first entry)						
XX							
DE	Human secreted protein gene 15 clone HSDES04.						
XX							
KW	Human; secreted protein; fusion protein; gene therapy; protein therapy;						
KW	diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia						
KW	developmental abnormality; foetal deficiency; blood; allergy; renal; ds;						
KW	immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;						
KW	inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;						
KW	cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus						
KW	osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;						
KW	endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.						
OS	Homo sapiens.						
XX							
PN	WO9856804-A1.						
PD							
PD	17-DEC-1998.						
XX							
PF	11-JUN-1998; 98WO-US12125.						
XX							
PR	02-OCT-1997; 97US-0061060.						
PR	13-JUN-1997; 97US-0049547.						
PR	13-JUN-1997; 97US-0049548.						
PR	13-JUN-1997; 97US-0049549.						
PR	13-JUN-1997; 97US-0049550.						
PR	13-JUN-1997; 97US-0049606.						
PR	13-JUN-1997; 97US-0049607.						
PR	13-JUN-1997; 97US-0049608.						
PR	13-JUN-1997; 97US-0049609.						
PR	13-JUN-1997; 97US-0049610.						
PR	13-JUN-1997; 97US-0049611.						
PR	13-JUN-1997; 97US-0050366.						
PR	13-JUN-1997; 97US-0050367.						
PR	13-JUN-1997; 97US-0050368.						
PR	13-JUN-1997; 97US-0050369.						
PR	13-JUN-1997; 97US-0050370.						
PR	13-JUN-1997; 97US-0050371.						
PR	13-JUN-1997; 97US-0050372.						
PR	13-JUN-1997; 97US-0050373.						
PR	13-JUN-1997; 97US-0050374.						
PR	13-JUN-1997; 97US-0050375.						
PR	13-JUN-1997; 97US-0050376.						
PR	13-JUN-1997; 97US-0050377.						
PR	13-JUN-1997; 97US-0050378.						
PR	13-JUN-1997; 97US-0050379.						
PR	13-JUN-1997; 97US-0050380.						
PR	13-JUN-1997; 97US-0050381.						
PR	13-JUN-1997; 97US-0050382.						
PR	13-JUN-1997; 97US-0050383.						
PR	13-JUN-1997; 97						

CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
XX
SQ Sequence 667 BP; 226 A; 143 C; 164 G; 134 T; 0 other;

Query Match 7.9%; Score 94; DB 21; Length 667;
Best Local Similarity 69.8%; Pred. No. 6.7e-06;
Matches 127; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

[illegible]

RESULT 8
AAC77124
ID AAC77124 standard; cDNA; 1722 BP.

AAC77124;

DT 08-FEB-2001 (first entry)

DE Human ORFX ORF2679 polynucleotide sequence SEQ ID NO: 5357.

KW Human: open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; anticholera; antifungal; antitubercumatic; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.

OS Homo sapiens.

PN WO200058473-A2.

05-OCT-2000.
PD

31-MAR-2000; 2000WO-US08621.

PR 31-MAR-1999: 99US-0127607.

PR 05-APR-1999; 99US-0127728;
02 APR 1999; 99US-0127030;

XX
XX
0000007, 0000008, 0000009, 0000010, 0000011, 0000012, 0000013, 0000014, 0000015, 0000016, 0000017, 0000018, 0000019, 0000020, 0000021, 0000022, 0000023, 0000024, 0000025, 0000026, 0000027, 0000028, 0000029, 0000030, 0000031, 0000032, 0000033, 0000034, 0000035, 0000036, 0000037, 0000038, 0000039, 0000040, 0000041, 0000042, 0000043, 0000044, 0000045, 0000046, 0000047, 0000048, 0000049, 0000050, 0000051, 0000052, 0000053, 0000054, 0000055, 0000056, 0000057, 0000058, 0000059, 0000060, 0000061, 0000062, 0000063, 0000064, 0000065, 0000066, 0000067, 0000068, 0000069, 0000070, 0000071, 0000072, 0000073, 0000074, 0000075, 0000076, 0000077, 0000078, 0000079, 0000080, 0000081, 0000082, 0000083, 0000084, 0000085, 0000086, 0000087, 0000088, 0000089, 0000090, 0000091, 0000092, 0000093, 0000094, 0000095, 0000096, 0000097, 0000098, 0000099, 0000100, 0000101, 0000102, 0000103, 0000104, 0000105, 0000106, 0000107, 0000108, 0000109, 0000110, 0000111, 0000112, 0000113, 0000114, 0000115, 0000116, 0000117, 0000118, 0000119, 0000120, 0000121, 0000122, 0000123, 0000124, 0000125, 0000126, 0000127, 0000128, 0000129, 0000130, 0000131, 0000132, 0000133, 0000134, 0000135, 0000136, 0000137, 0000138, 0000139, 0000140, 0000141, 0000142, 0000143, 0000144, 0000145, 0000146, 0000147, 0000148, 0000149, 0000150, 0000151, 0000152, 0000153, 0000154, 0000155, 0000156, 0000157, 0000158, 0000159, 0000160, 0000161, 0000162, 0000163, 0000164, 0000165, 0000166, 0000167, 0000168, 0000169, 0000170, 0000171, 0000172, 0000173, 0000174, 0000175, 0000176, 0000177, 0000178, 0000179, 0000180, 0000181, 0000182, 0000183, 0000184, 0000185, 0000186, 0000187, 0000188, 0000189, 0000190, 0000191, 0000192, 0000193, 0000194, 0000195, 0000196, 0000197, 0000198, 0000199, 0000200, 0000201, 0000202, 0000203, 0000204, 0000205, 0000206, 0000207, 0000208, 0000209, 0000210, 0000211, 0000212, 0000213, 0000214, 0000215, 0000216, 0000217, 0000218, 0000219, 0000220, 0000221, 0000222, 0000223, 0000224, 0000225, 0000226, 0000227, 0000228, 0000229, 0000230, 0000231, 0000232, 0000233, 0000234, 0000235, 0000236, 0000237, 0000238, 0000239, 0000240, 0000241, 0000242, 0000243, 0000244, 0000245, 0000246, 0000247, 0000248, 0000249, 0000250, 0000251, 0000252, 0000253, 0000254, 0000255, 0000256, 0000257, 0000258, 0000259, 0000260, 0000261, 0000262, 0000263, 0000264, 0000265, 0000266, 0000267, 0000268, 0000269, 0000270, 0000271, 0000272, 0000273, 0000274, 0000275, 0000276, 0000277, 0000278, 0000279, 0000280, 0000281, 0000282, 0000283, 0000284, 0000285, 0000286, 0000287, 0000288, 0000289, 0000290, 0000291, 0000292, 0000293, 0000294, 0000295, 0000296, 0000297, 0000298, 0000299, 0000300, 0000301, 0000302, 0000303, 0000304, 0000305, 0000306, 0000307, 0000308, 0000309, 0000310, 0000311, 0000312, 0000313, 0000314, 0000315, 0000316, 0000317, 0000318, 0000319, 0000320, 0000321, 0000322, 0000323, 0000324, 0000325, 0000326, 0000327, 0000328, 0000329, 0000330, 0000331, 0000332, 0000333, 0000334, 0000335, 0000336, 0000337, 0000338, 0000339, 0000340, 0000341, 0000342, 0000343, 0000344, 0000345, 0000346, 0000347, 0000348, 0000349, 0000350, 0000351, 0000352, 0000353, 0000354, 0000355, 0000356, 0000357, 0000358, 0000359, 0000360, 0000361, 0000362, 0000363, 0000364, 0000365, 0000366, 0000367, 0000368, 0000369, 0000370, 0000371, 0000372, 0000373, 0000374, 0000375, 0000376, 0000377, 0000378, 0000379, 0000380, 0000381, 0000382, 0000383, 0000384, 0000385, 0000386, 0000387, 0000388, 0000389, 0000390, 0000391, 0000392, 0000393, 0000394, 0000395, 0000396, 0000397, 0000398, 0000399, 0000400, 0000401, 0000402, 0000403, 0000404, 0000405, 0000406, 0000407, 0000408, 0000409, 0000410, 0000411, 0000412, 0000413, 0000414, 0000415, 0000416, 0000417, 0000418, 0000419, 0000420, 0000421, 0000422, 0000423, 0000424, 0000425, 0000426, 0000427, 0000428, 0000429, 0000430, 0000431, 0000432, 0000433, 0000434, 0000435, 0000436, 0000437, 0000438, 0000439, 0000440, 0000441, 0000442, 0000443, 0000444, 0000445, 0000446, 0000447, 0000448, 0000449, 0000450, 0000451, 0000452, 0000453, 0000454, 0000455, 0000456, 0000457, 0000458, 0000459, 0000460, 0000XX
(CORA)) CONAGEN CORP.[illegible]

DR P-PSDB; AAB42915.

PT-5 Novel nucleic acids and peptides derived from open reading frame X₁

useful for treating e.g. cancers, proliferative disorders,
neurodegenerative disorders and cardiovascular disease -
Claim 5; Page 4535-4536; 5507pp; English.

AC74446 to AAC77606 encode the proteins given in ABA40237 to ABA43397,
which represent the human ORFX open reading frames 1 to 3161. The ORFX
sequences have activities such as: cytostatic; hepatotropic; vulvular;
antiproliferative; antiparkinsonian; nootropic; neuroprotective;
osteopathic; anticoagulant; antiarthritic; immunosuppressive;
immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
antidiabetic; hypotensive; dermatological; immunosuppressive;
antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
antithyroid; and antianemic. The sequences can be used for determining
the presence of or predisposition to, or preventing or treating
pathological conditions associated with an ORFX-associated disorder. The
nucleic acids can be used to express ORFX proteins in gene therapy
vectors. The proteins and nucleic acids may be used to treat cancers,
proliferative disorders, neurodegenerative disorders, osteoarthritis,
graft vs host disease, cardiovascular disease, diabetes mellitus,
hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
bacterial or fungal infection, malaria, autoimmune disorders, asthma,
allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
nocturnal haemoglobinuria, antiinflammatory disease; to enhance
coagulation; to inhibit thrombosis; and as a contraceptive.

Sequence 1722 BP; 484 A; 421 C; 427 G; 390 T; 0 other;

Query Match 7.8%; Score 92.8; DB 21; Length 1722;
Best Local Similarity 82.8%; Pred. No. 1e-05;
Matches 106; Conservative 0; Mismatches 22; Indels 0; Gaps 0

Qy	1062	gccaggccacagagcgagattccatctc	caaaaaataataataataataataaaa	1122a
Db	1595	gctgagccctctgctataggaaaaaaa	aaaaaaaaaaaaaaaaaaaaaaaaaaaaa	1654a
Qy	1122	tataaaaaataaaaaaaa	aaaaaaaaaaaaaaaaaaaaaaaaaaaaa	1187a
Db	1655	aaaaaaataaaaaaaa	aaaaaaaaaaaaaaaaaaaaaaaaaaaaa	1714a

Qy 1182 aaaaaaa 1189
 | | | | |
Db 1715 aaaaaaa 1722

RESULT 9
AAC59838
ID AAC59838 standard: DNA: 1529 BP.

AA	AAC59838;
AC	26-JAN-2001
XX	(first entry)
DT	Human secreted protein encoding DNA clone vc69 1.
XX	
DE	

Secreted protein; human; autoimmune disorder; multiple sclerosis; ulcer;
 KW
 Systemic lupus erythematosus; rheumatoid arthritis; anaemia; stroke;
 KW
 Haematopoiesis regulation; tissue growth; wound healing; haemophilia;
 KW
 Alzheimer's disease; Parkinson's disease; Shy-drager syndrome; cancer;
 KW
 Contraceptive; infection; growth inhibition; hyperproliferative disorder
 KW
 Psoriasis; ds.

OS Homo sapiens.

PN WO200055375-A1.

PD 21-SEP-2000.

PF 17-MAR-2000; 2000WO-US07285.

PR 17-MAR-1999; 99US-0124808.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 6, 2001, 21:05:44 ; Search time 2566.06 Seconds
(without alignments)
7173.106 Million cell updates/sec

Title: US-09-391-861-1
Perfect score: 1190
Sequence: 1 gaggatccagccgaagagg.....aaaaaaaaaaaaaaaaaaaa 1190

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_v1:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_v11:*
59: gb_v12:*
60: gb_vt1:*
61: gb_vt2:*
62: gb_vt3:*
63: gb_vt4:*
64: gb_vt5:*
65: gb_vt6:*
66: gb_vt7:*
67: gb_vt8:*
68: gb_vt9:*
69: gb_vt10:*
70: gb_vt11:*
71: gb_vt12:*
72: gb_vt13:*
73: gb_vt14:*
74: gb_vt15:*
75: gb_vt16:*
76: gb_vt17:*
77: gb_vt18:*
78: gb_vt19:*
79: gb_vt20:*
80: gb_vt21:*
81: gb_vt22:*
82: gb_vt23:*
83: gb_vt24:*
84: gb_vt25:*
85: gb_vt26:*
86: gb_vt27:*
87: gb_vt28:*
88: gb_vt29:*
89: gb_vt30:*
90: gb_vt31:*
91: gb_vt32:*
92: gb_vt33:*
93: gb_vt34:*
94: gb_vt35:*
95: gb_vt36:*
96: gb_vt37:*
97: gb_vt38:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Match	Length	DB	ID	Description
1	628.4	52.8	630	85	AB021975	Homo sapi	
C 2	575	48.3	37402	87	AC009002	Homo sapi	
C 3	563	47.3	110000	60	AC008749_0	Homo sapi	
C 4	563	47.3	110000	60	AC008749_1	Continuation (2 of	
C 5	556.2	46.7	4033	85	AB006136	Homo sapi	
C 6	495	41.6	514	10	AX097639	Sequence	
C 7	435.8	36.6	227949	70	AC026803	Homo sapi	
C 8	412.6	34.7	179538	68	AC024740	Homo sapi	

ALIGNMENTS

```

RESULT 1
LOCUS AB0211975 630 bp mRNA PRI
DEFINITION Homo sapiens mRNA for FGF-21, complete cds.
ACCESSION AB021975
VERSION AB021975.1 GI:9049444
KEYWORDS new FGF.
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Nishimura,T., Nakatake,Y., Konishi,M. and Itoh,N.
TITLE Identification of a novel FGF, FGF-21, preferentially expressed in
the liver(1)
MEDLINE Biochim. Biophys. Acta 1492 (1), 203-206 (2000)
REFERENCE 20461777
AUTHORS Itoh,N.
TITLE Direct Submission
JOURNAL Submitted (29-DEC-1998) to the DDBJ/EMBL/GenBank databases
Nobuyuki Itoh, Kyoto University, Graduate School of Pharmaceutical
Sciences, Department of Genetic Biochemistry; Yoshida-Shimoadachi,
Sakyo, Kyoto, Kyoto 606-8501, Japan
(E-mail: itohnobu@pharm.kyoto-u.ac.jp. Tel:81-75-753-4540,
Fax:81-75-753-4600)
COMMENT Sequence updated (17-Mar-1999).
FEATURES
location/Qualifiers
1..630
/organism="Homo sapiens"
/db_xref="taxon:9606"
source

```


6233 7544: contig of 1312 bp in length
7545 gap of unknown length
7645 contig of 1248 bp in length
8892 gap of unknown length
8893 contig of 1304 bp in length
8993 10296: gap of unknown length
10297 10396: gap of unknown length
10397 11725: contig of 1329 bp in length
11726 11825: gap of unknown length
11826 13734: contig of 1909 bp in length
13735 13834: gap of unknown length
13835 15056: contig of 1222 bp in length
15057 15156: gap of unknown length
15157 16358: contig of 1202 bp in length
16359 16458: gap of unknown length
16459 17875: contig of 1417 bp in length
17876 17975: gap of unknown length
17976 19527: contig of 1552 bp in length
19527 19627: gap of unknown length
19628 21424: contig of 1797 bp in length
21425 21524: gap of unknown length
21525 22793: contig of 1269 bp in length
22794 22893: gap of unknown length
22894 23905: contig of 1012 bp in length
23906 24005: gap of unknown length
24006 25440: contig of 1435 bp in length
25441 25540: gap of unknown length
25541 26895: contig of 1355 bp in length
26896 26995: gap of unknown length
26996 28178: contig of 1183 bp in length
28179 28278: gap of unknown length
28279 29887: contig of 1609 bp in length
29888 29987: gap of unknown length
29989 31869: contig of 1882 bp in length
31870 33535: gap of unknown length
33536 33635: contig of 1566 bp in length
33636 34963: contig of 1328 bp in length
34964 35063: gap of unknown length
35064 36881: contig of 1818 bp in length
36882 38148: gap of unknown length
38149 38248: contig of 1167 bp in length
38249 39889: contig of 1641 bp in length
39890 41963: gap of unknown length
41964 42063: contig of 1974 bp in length
42064 43733: gap of unknown length
43734 43833: gap of unknown length
43834 45414: contig of 1581 bp in length
45415 45514: gap of unknown length
45516 47046: contig of 1532 bp in length
47047 47146: gap of unknown length
47147 48965: contig of 1819 bp in length
48966 49065: gap of unknown length
49066 51369: contig of 2304 bp in length
51370 51469: gap of unknown length
51470 53138: contig of 1659 bp in length
53139 53228: gap of unknown length
53229 54480: contig of 1252 bp in length
54481 54580: gap of unknown length
54581 56485: contig of 1905 bp in length
56486 56585: gap of unknown length
56586 58664: contig of 2079 bp in length
58665 58764: gap of unknown length
58765 59959: contig of 1195 bp in length
59960 60059: gap of unknown length
60060 62303: contig of 2244 bp in length
62304 62403: gap of unknown length
62404 64304: contig of 1801 bp in length
64305 64306: gap of unknown length
64307 66196: contig of 1892 bp in length
66197 66296: gap of unknown length
66297 67878: contig of 1582 bp in length

67879 67978: gap of unknown length
67979 69965: contig of 1987 bp in length
69966 70065: gap of unknown length
70066 71633: contig of 1568 bp in length
71634 71733: gap of unknown length
71734 73978: contig of 2245 bp in length
73979 74078: gap of unknown length
74079 76073: contig of 1995 bp in length
76074 76173: gap of unknown length
76174 78469: contig of 2296 bp in length
78470 78569: gap of unknown length
78570 81250: contig of 2681 bp in length
81251 83762: gap of unknown length
83763 83862: gap of unknown length
83863 85989: contig of 2127 bp in length
85990 86089: gap of unknown length
86090 88156: contig of 2067 bp in length
88157 88256: gap of unknown length
88257 90418: contig of 2062 bp in length
90419 92289: gap of unknown length
92290 92390: contig of 2785 bp in length
92391 95174: gap of unknown length
95175 95274: gap of unknown length
95275 98333: contig of 3059 bp in length
98334 98433: gap of unknown length
98434 99975: contig of 1542 bp in length
99976 100075: gap of unknown length
100076 101499: contig of 1424 bp in length
101500 101599: gap of unknown length
101600 103848: contig of 2249 bp in length
103849 103948: gap of unknown length
103949 106121: contig of 2179 bp in length
106122 106227: gap of unknown length
106228 108402: contig of 2175 bp in length
108403 108502: gap of unknown length
108503 110948: contig of 2446 bp in length
110949 111048: gap of unknown length
111049 113905: contig of 2857 bp in length
113906 114005: gap of unknown length
114006 117256: contig of 3251 bp in length
117257 117356: gap of unknown length
117357 119135: contig of 1779 bp in length
119136 119235: gap of unknown length
119236 121325: contig of 2090 bp in length
121326 121425: gap of unknown length
121426 123954: contig of 2529 bp in length
123955 124054: gap of unknown length
124055 126851: contig of 2797 bp in length
126852 126951: gap of unknown length
126952 129652: contig of 2701 bp in length
129653 129752: gap of unknown length
129753 132233: contig of 2481 bp in length
132234 132333: gap of unknown length
132334 135268: contig of 2935 bp in length
135269 135368: gap of unknown length
135369 137746: contig of 2378 bp in length
137747 137846: gap of unknown length
137847 141549: contig of 3703 bp in length
141550 141649: gap of unknown length
141650 144391: contig of 2742 bp in length
144392 144491: gap of unknown length
144492 147414: contig of 2923 bp in length

Query Match 47.3%; Score 563; DB 60; Length 110000;

Best Local Similarity 97.3%; Pred. No.1e-80; Indels 1; Gaps 1;
Matches 583; Conservative 0; Mismatches 15;

QY 478 tcgctccactttgaccctgagcctgagcctccgagcttccggagctcttcgagacgatac 537
DB 109757 TAGCTCCACTTTGACCCTGAGGCGCTGCAGCTCCGGAGCTGCTTCTTGAGACGATAC 109698

```
QY 538 aatgtttaccagtcgaagcccaagcctcccgctcccgctgacatgccaaggaacaagtcccca 597
Db 109697 AATGTTTACCAAGTCCGAAGCCACCGGCTCCCGCTGCACCTGCCAGGAACAAGTCCCCA 109638
QY 598 caccggaccctgcaccccgaggaagcagctcgcttccctgcaactacacagcctgcacccc 657
Db 109637 CACCGGACCCCTGCACCCCGAGGACCAAGCTGCTTCCCTGCACCTACCAAGCCCTGCCCCC 109578
QY 658 gcaccccccagagccaccccggaatcctcgccccccagccccccgagtgctgggctcctcgag 717
Db 109577 GCACCCCGGAGCCACCCGGAATCCTGGCCCCCAGCCGCCCGATGTGGGCTCCTCGGAC 109518
QY 718 cctctgagcatggtggaccctccagggccggaagccccagctacgcttccctgaaagccag 777
Db 109517 CCTCTGAGCATGCTGGGACCTTCCAGGCGCGAAGCCCGACGCTACGCTTCTCCTGAAGCCAG 109458
QY 778 aggtgtttactatgacatcctcttattattattaggtattattattattatttttttttt 837
Db 109457 AGGCTGTTTACTATGACATCTCCTCTTTATTATTAGGTTATTATTATTATTATTATT 109398
QY 838 tatttttcttacttgagataataaagagttccagagagagataaagatgagcatgtgtga 897
Db 109397 TATTTTTTCTTACTTGAGATAATAAAGAGTTCACAGAGGAGGATAAAGATGAGCATGTGTGA 109338
QY 898 gtgtctgaggaagacaatggcagctgtttgtctccctggcccggaacaatccccctctac 957
Db 109337 GTGTCTGAGGGAAGACATGGCAGCTGTTTGTCTTCTGGCCCGGACAAACCCCTCTAC 109278
QY 958 acctccctcacgtggtccgagggctggtgctccacacagggcctcaacttttttttttt 1017
Db 109277 ACCTCCCTCACGTGTCGAGGGTCTGCTGCTTCCACCTGGGCTCACATTTTCTTTTTC 109218
QY 1018 ttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt 1076
Db 109217 TTTTCTTTT-TTTTTTTTGAGACGGAGTCTGCTGTGTGACCCAGCGTGGAGTGAGTG 109160

RESULT 4
AC008749_1/c
WPCOMMENT
Sequence split into 5 fragments LOCUS AC008749 Accession AC008749
Fragment Name Begin End
AC008749_0 1 110000
AC008749_1 100001 210000
AC008749_2 200001 310000
AC008749_3 300001 410000
AC008749_4 400001 467420
Continuation (2 of 5) of AC008749 from base 100001 (AC008749 Homo sapiens chromosome 19

Query Match 47.3%; Score 563; DB 60; Length 110000;
Best Local Similarity 97.3%; Pred. No. 1e-80;
Matches 583; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 478 tgcgtccacttgacacctgagcctgcagcttcgaggtctcttgagacgataac 537
Db 9757 TAGCTCCACTTTGACCTTGAGCCCTGCAGCTTCGGGAGCTGCTTCTTGAGGACGGATAC 9698
QY 538 aatgtttaccagtcgaagcccaagcctcccgctgcaactgccaggaacaagtcccca 597
Db 9697 AATGTTTACAGTCCGAGAGCCACCGGCTCCCGCTGCACCTGCCAGGGAACAAGTCCCCA 9638
QY 598 caccggaccctgcaccccgaggaagaccagctcgcttccctgcaactacacagggcctgcacccc 657
Db 9637 CACCGGACCCCTGCACCCCGAGGACCAAGCTGCTTCTTCTGCTTCCACTACCAAGGCTGCCCCC 9578
QY 658 gcaccccccagagccaccccggaatcctgccccccagccccccagctggtggtcctcgag 717
Db 9577 GCACCCCGGAGCCACCCGGAATCCTGGCCCCCAGCCGCCCGATGTGGGCTCCTCGGAC 9518
QY 718 cctctgagcatggtggaccctccagggccgaagccccagctacgcttccctgaagccag 777
Db 9517 CCTCTGAGCATGCTGGGACCTTCCAGGCGCGAAGCCCGACGCTACGCTTCTCCTGAAGCCAG 9458
```

```
QY 778 aggtgtttactatgacatcctcctcttattattattaggtattattattattattttttt 837
Db 9457 AGGCTGTTTACTATGACATCTCCTCTTTATTATTATTAGGTTATTATTATTATTATT 9398
QY 838 tatttttcttacttgagataataaagagttccagagagataaagatgagcatgtgtga 897
Db 9397 TATTTTTTCTTACTTGAGATAATAAAGAGTTCAGAGGAGGATAAAGATGAGCATGTGTGA 9338
QY 898 gtgtctgaggaagacaatggcagctgtttgtctccctggcccggaacaatccccctctac 957
Db 9337 GTGTCTGAGGGAAGACATGGCAGCTGTTTGTCTTCTTGGCCCGGACAAACCCCTCTAC 9278
QY 958 acctccctcacgtggtccgagggctcggttccacacagggcctcaacttttttttttt 1017
Db 9277 ACCTCCCTCACGTGTCGAGGGTCTGCTGCTTCCACCTGGGCTCACATTTTCTTTTTC 9218
QY 1018 ttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt 1076
Db 9217 TTTTCTTTT-TTTTTTTTGAGACGGAGTCTGCTGTGTGACCCAGCGTGGAGTGAGTG 9160

RESULT 5
AB006136/c
LOCUS
DEFINITION Homo sapiens gene for alpha 1,2-fucosyltransferase, 5' flanking
region and partial cds.
ACCESSION AB006136
VERSION AB006136.1 GI:3242453
KEYWORDS FUT1; alpha (1, 2) fucosyltransferase.
SOURCE Homo sapiens peripheral leukocytes DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Koda,Y., Soejima,M. and Kimura,H.
TITLE Changing transcription start sites in H-type
alpha(1,2)fucosyltransferase gene (FUT1) during differentiation of
the human erythroid lineage
Eur. J. Biochem. 256 (2), 379-387 (1998)
98430978
REFERENCE 2 (bases 1 to 4033)
AUTHORS Koda,Y.
TITLE Direct Submission
JOURNAL Submitted (04-AUG-1997) to the DDBJ/EMBL/GenBank databases. Yoshiro
Koda, Kurume University, School of Medicine, Department of Forensic
Medicine, Asahimachi 67, Kurume, Fukuoka 830-0011, Japan
(E-mail:ykoda@med.kurume-u.ac.jp, Tel:81-942-31-7554,
Fax:81-942-31-7700)
COMMENT On Jun 20, 1998 this sequence version replaced gi:2317262.
Sequence updated (16-Jun-1998).
FEATURES
Location/Qualifiers
source
1..4033
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.3"
/tissue_type="peripheral leukocytes"
promoter
1..510
exon
511..578
/gene="FUT1"
/number=1
511..578
/gene="FUT1"
511..>578
/gene="FUT1"
/standard_name="alpha(1,2)fucosyltransferase"
/codon_start=1
/protein_id="BAA28952.1"
/db_xref="GI:3242453"
/translation="MGFHHVQGAGLELLTSGDLPAMV"

BASE COUNT 877 a 1242 c 1099 g 815 t
```

ORIGIN

Query Match	46.7%;	Score 556.2;	DB 85;	Length 4033;
Best Local Similarity	96.7%;	Pred. No. 1.3e-79;		
Matches 579; Conservative	0;	Mismatches 18;	Indels 2;	Gaps 1;

478	tcgtccacattgaacctgaggcctgcagctctccgggaagctgctctcttgaaggacgatac	537
479		
480		
481		
482		
483		
484		
485		
486		
487		
488		
489		
490		
491		
492		
493		
494		
495		
496		
497		
498		
499		
500		
501		
502		
503		
504		
505		
506		
507		
508		
509		
510		
511		
512		
513		
514		
515		
516		
517		
518		
519		
520		
521		
522		
523		
524		
525		
526		
527		
528		
529		
530		
531		
532		
533		
534		
535		
536		
537		
538		
539		
540		
541		
542		
543		
544		
545		
546		
547		
548		
549		
550		
551		
552		
553		
554		
555		
556		
557		
558		
559		
560		
561		
562		
563		
564		
565		
566		
567		
568		
569		
570		
571		
572		
573		
574		
575		
576		
577		
578		
579		
580		
581		
582		
583		
584		
585		
586		
587		
588		
589		
590		
591		
592		
593		
594		
595		
596		
597		
598		
599		
600		
601		
602		
603		
604		
605		
606		
607		
608		
609		
610		
611		
612		
613		
614		
615		
616		
617		
618		
619		
620		
621		
622		
623		
624		
625		
626		
627		
628		
629		
630		
631		
632		
633		
634		
635		
636		
637		
638		
639		
64		

RESULT	6
AX097639	
LOCUS	514 bp DNA
DEFINITION	Sequence 1 from Patent WO0118209.
ACCESSION	AX097639
VERSION	AX097639.1 GI:13514271
KEYWORDS	.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 514)
AUTHORS	Shimkets,R.A., Vernet,C., Burgess,C., Fernandes,E., Taupier,R.J., Quinn,K.E., Spyteck,K.A., Rastelli,L. and Herrmann,J.L.
TITLE	Fibroblast growth factor polypeptide and nucleic acids encoding same
JOURNAL	Patent: WO 0118209-A 1 15-MAR-2001;
FEATURES	Curagen Corporation (US) Location/Qualifiers 1..514 /organism="Homo saplens" /db_xref="taxon:9606" 52..513 /note="unnamed protein product"
source	/cdote start=1
CDS	

/protein_id="CAC35432.1"
 /db_xref="GI:13514372"
 /translation="MDSDETFEHSGLVFWVLGILLCAOAHPIPDSSPLPFGGV
 QRYYLYTDQAEALRETRGDTVGGRADSPSLQLKALKPGVYQLIGVXTSL
 CPQDAGLYSLDFDEACFREDLGGDGYNYQSEAHGLFHLPLGQRRL"
 103 a 157 c 155 g 99 t
 BASE COUNT
 ORIGIN

Query Match	41.68;	Score 495;	DB 10;	Length 514;
Best Local Similarity	100.0%;	Pred. No. 7.5e-70;		
Matches 495;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	91	caqcttctccagatccacaccccgagatcacctcgaggaccgagccattgatggactcg	150	
DB	1	CAGCTTCTCGAGCTCACACCCCGAGATCACCTGAGGACCGAGCATTTGATGGACTCG	60	
QY	151	gacgagacgggttcgagcactcaggactgtgggtttctgtcgtcgtggtttctcgtg	210	
DB	61	GACGAGACGGGTTCCGAGCACTCAGGACTGTGGGTTTCTGTGCTGGCTGGTCTTGCTG	120	
QY	211	ggagcctgccaggcacaccccatccctgactccagtcctcctcgaattcgggggcca	270	
DB	121	GGAGCCTGCCAGGCACACCCCATCCCTGACTCCAGTCCTCTCTGCAATTCGGGGGCCNA	180	
QY	271	gtccggcagcggtacctctacacagatgatgccagacagagaagccacactggagatc	330	
DB	181	GTCCGGCAGCGTACTCTACAGATGATGCCGAGCAGACAGAACCCACCTGGAGATC	240	
QY	331	agggagatgggacggtggggggcgctgctgaccagagccccaaagtctcctgacgtg	390	
DB	241	AGSGAGATGGACGGTGGGGGGCGCTGCTGACCAGAGCCCGAAAAGTCTCTCTGCAGCTG	300	
QY	391	aaagcctgaagccggggagttattcaaatcttgggagtcgaagacatccaggttcctgtgc	450	
DB	301	AAAGCCTTGAAGCCGGGAGTTATTCAAACTCTGGGAGTCGAAGACATCCAGGTTCCCTGTGC	360	
QY	451	caaggccagatggggccctgtatggatcgctcgaattgaccctgagggcctgcagcttc	510	
DB	361	CAGCGGCCAGATGGGGCCCTGTATGGATCGTCCACTTTGACCTTGAGGCTTGACCTTC	420	
QY	511	cgggagctgctcttgaggacggatacaattgtaccagttccgaagccacggcctcccg	570	
DB	421	CGGAGCTGCTTCTTGAGCAGGATACAAATGTTTACCAGTCCGAAGCCACGSCCTCCGG	480	
QY	571	ctgcacctgccaggg	585	
DB	481	CTGCACTGCCAGG	495	

[illegible]

Consensus quality: 173830 bases at least Q40
Consensus quality: 175557 bases at least Q30
Consensus quality: 176446 bases at least Q20
Insert size: 242000; agarose-fp
Insert size: 177938; sum-of-continigs
Quality coverage: 5.32 in Q20 bases; agarose-fp
Quality coverage: 6.01 in Q20 bases; sum-of-continigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

* 1 1801: contig of 1801 bp in length
* 1802 1901: gap of unknown length
* 1902 4641: contig of 2740 bp in length
* 4642 4741: gap of unknown length
* 4742 9679: contig of 4938 bp in length
* 9680 9779: gap of unknown length
* 9780 13392: contig of 3613 bp in length
* 13393 13492: gap of unknown length
* 13493 17164: contig of 3672 bp in length
* 17165 17264: gap of unknown length
* 17265 21510: contig of 4246 bp in length
* 21511 21610: gap of unknown length
* 21611 27089: contig of 5479 bp in length
* 27090 27189: gap of unknown length
* 27190 35636: contig of 8447 bp in length
* 35637 35736: gap of unknown length
* 35737 45537: contig of 9801 bp in length
* 45538 56537: gap of unknown length
* 56539 56451: contig of 10814 bp in length
* 56452 66802: contig of 10251 bp in length
* 66803 66902: gap of unknown length
* 66903 79118: contig of 12216 bp in length
* 79119 79218: gap of unknown length
* 79219 94355: contig of 15137 bp in length
* 94356 94455: gap of unknown length
* 94456 111274: contig of 16819 bp in length
* 111275 111374: gap of unknown length
* 111375 130600: contig of 19226 bp in length
* 130601 130700: gap of unknown length
* 130701 150278: contig of 19578 bp in length
* 150279 150378: gap of unknown length
* 150379 179538: contig of 29160 bp in length.

FEATURES Location/Qualifiers

1..179538
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="RP11-801D6"
1..1801

misc_feature /note="assembly_name:Contig13"
misc_feature 1902..4641
misc_feature /note="assembly_name:Contig14"
misc_feature 4742..9679
misc_feature /note="assembly_name:Contig15"
misc_feature 9780..13392
misc_feature /note="assembly_name:Contig16"
misc_feature 13493..17164
misc_feature /note="assembly_name:Contig17"
misc_feature 17265..21510
misc_feature /note="assembly_name:Contig18"
misc_feature 21611..27089
misc_feature /note="assembly_name:Contig19"
misc_feature 27190..35636
misc_feature /note="assembly_name:Contig20"
misc_feature 35737..45537
misc_feature /note="assembly_name:Contig21"

misc_feature 45638..56451
/note="assembly_name:Contig22"
misc_feature 56552..66802
/note="assembly_name:Contig23"
misc_feature 66903..79118
/note="assembly_name:Contig24"
misc_feature 79219..94355
/note="assembly_name:Contig25"
misc_feature 94456..111274
/note="assembly_name:Contig26"
misc_feature 111375..130600
/note="assembly_name:Contig27"
misc_feature 130701..150278
/note="assembly_name:Contig28
vector_side:left"
misc_feature 150379..179538
/note="assembly_name:Contig29"
BASE COUNT 42230 a 47023 c 46289 g 42392 t 1604 others
ORIGIN

Query Match 34.7%; Score 412.6; DB 68; Length 179538;
Best Local Similarity 80.0%; Pred. No. 8.9e-57;
Matches 415; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

Qy 478 tcgctccactttgacctgagcctgagctgcagctccggagctgctctcttggagcgatcac 537
Db 130186 TAGCTCCACCTTGACCTGAGCCCTGAGCTTCCGGAGCTGCTTCTTGGAGGACGATAC 130245
Qy 538 aatgtttaccagtcgagccacgagcctccctccctgcagctgcagctgcagggagcaagtcacca 597
Db 130246 AATGTTTACCAGTCCGAAGCCACGGGCTCCCGCTGCACCTGCAGGGAACAAGTCCCA 130305
Qy 598 caccgggacctgcaccccgagcagcagctgcgcttcttgcacaccagcctgcacccccc 657
Db 130306 CACCGGGACCTGCACCCCGAGGACCACTCGTCTCGCCACTACCAAGCTGCCCCCC 130365
Qy 658 qcaccccgagcagccacccggaatcctgagcccccagcccccagctgtgggctcctcggac 717
Db 130366 GCACCCCGGAGCACCCCGGAATCTGTGCCCCCGCCAGCCCCCGATGTGGCTCTCGGAC 130425
Qy 718 cctctgagcatggtgggaccttccagggccgagcccccagctacgcttctcctgaagccag 777
Db 130426 CCTCTGAGCATGFTGGACCTTCCAGGGCCGAAGCCAGCTATGCTTCTCTGAAGCCAG 130485
Qy 778 aggtgttactatgacatcctcctcttattattattattattattattatttttttttttttttt 837
Db 130486 AGGCTGTTTACTATGACATCTCTCTTTATTTATTTATTTATTTATTTATTTATTTT 130545
Qy 838 tattttcttacttgagataaagagttccagagagagagataagatgacatgtgtga 897
Db 130546 TATTTTCTTACTTGAGATAATAAAGAGTTCCAGAGGAGGATAAGATGAGCATGNNNNN 130605
Qy 898 gtctctgaggaagacatggcagctgtttgtctccttggcccgagacatccccctctac 957
Db 130606 NNN 130665
Qy 958 acctccccctacgtggtccgagggctcctgctccacct 996
Db 130666 NNN 130704

RESULT 9

AB025718
LOCUS AB025718 633 bp mRNA ROD 11-JUL-2000
DEFINITION Mus musculus mRNA for FGF-21, complete cds.
ACCESSION AB025718
VERSION AB025718.1 GI:9049446
KEYWORDS FGF-21.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (sites)
Nishimura,T., Nakatake,Y., Konishi,M. and Itoh,N.
Identification of a novel FGF, FGF-21, preferentially expressed in
the liver(1)
Biochim. Biophys. Acta 1492 (1), 203-206 (2000)
20461777
2 (bases 1 to 633)
Itoh,N.
Direct Submission
Submitted (02-APR-1999) to the DDBJ/EMBL/GenBank databases.
Nobuyuki Itoh, Kyoto University, Graduate School of Pharmaceutical
Sciences, Department of Genetic Biochemistry; Yoshida-Shimoedachi,
Sakyo, Kyoto, Kyoto 606-8501, Japan
(E-mail:iitohnobuepharm.kyoto-u.ac.jp, Tel:81-75-753-4540,
Fax:81-75-753-4600)

FEATURES
source
1..633
/organism="Mus musculus"
/db_xref="taxon:10090"
1..633
/gene="FGF-21"
1..633
/gene="FGF-21"
/codon_start=1
/product="FGF-21"
/protein_id="BAA99416.1"
/db_xref="GI:9049447"
/translation="MEWMSRVCTGLVRLLLAVFLLVGYOAYPIPDSPILQFGGQ
VQRVLYTDDQDTEAHLREDCYVGAHRSPELLELKALKPVIQLGVKASRF
LCQDPGALIGSFHPDECFSELLEDGVNYQEAGHLPLRLPKDKSPNDATSW
GGVPRLFGLIHPEODQAQFIPEPPDVGSSDPLSMVEFPLOGRSPSYAS"

BASE COUNT
136 a 198 c 173 g 126 t

ORIGIN

Query Match	33.8%; Score 401.8; DB 94; Length 633;
Best Local Similarity	80.7%; Pred.No.5.1e-55;
Matches 469; Conservative	0; Mismatches 112; Indels 0; Gaps 0;

```

QY 191 tgctggctgtcttctgctggagcctgccagcacaccccatccctgcactccagtctc 250
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 53 TGTGGCTGTCTCCTGCTGGGGTCTACCAAGCATACCCTATCCCTGTATCCAGCCCC 112
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 251 tcctgcaattcggggcgaagtcgcgcagcgtacctcttacagatgatgccagcaga 310
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 113 TCCTCCAGTTTGSGGTC AAGTCGGCGAGAGGTACCTCTACACAGATGACACCAACA 172
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 311 cagaagcccaccctggagatcaggagagatggagcgtggtggggcgctgtgcacagacc 370
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 173 CTGAAGCCCACCTGGAGATCAGGAGGATGGAACAGTGTGTAGCGCACACCGCAGTC 232
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 371 ccgaagatccctgcagctgaagccttgagccggagagattattcaaatctggagatca 430
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 233 CAGAAAGTCCTCTGGAGCTCAAAGCCTTGAAGCCAGGGGTCAATTCAAATCTCGGTGTC 292
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 431 agacatcaggttcctgtgcccagcgccagatgggcccctgtgtggatcgtccactttg 490
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 293 AAGCCCTAGGTTCTTTGCCAACGCCACGATGAGGCTCTCTATGGATCCGCTCACATTG 352
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 491 acctgaggcctgcagctccggagctgtctcttggagcagatcaaatgtttaccagt 550
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 353 ATCCTGAGGCCTGCAGCTTCAGAGAACTGCTGCTGGAGGACGGTTACAATGTGTACCA 412
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 551 ccgaagccccagcctccgctgcacctgcagggaaacaagtccccaacacggaccctg 610
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 413 CTGAAGCCCATGCGCTGCCCTCGCTGTGCTCAGAGGACTCCCCAAGCAGGATGCAA 472
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 611 cacccgagagaccagctcgcttctccactaccagcctgcccccgaccccccgagc 670
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 473 CATCTCTGGGACCTGTGCGCTTCTGTGCCATGCCAGGCGCTGTCTCCACGAGCCCCAAG 532
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 671 caccggaaatcctggccccccagccccccaatgtgggctctcctcgagccctctagaatgg 730
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db	533	AAGCAGGATTCTGCCCCAGAGAGCCCCAGATGGGCTCTCTGACCCCTGAGCATGG	592
Qy	731	tgggaccttcacagggccgaagccccagctacgcttctctga	771
Db	593	TAGAGCCTTTACAGGGCGGAAGCCCCAGCTATGCGTCTCTGA	633
RESULT 10			
AF214655/c			
LOCUS	AF214655	9257 bp	DNA
DEFINITION	Mus musculus alpha(1,2)fucosyltransferase FUT1 (Fut1) gene, complete cds.	23-MAR-2000	
ACCESSION	AF214655		
VERSION	AF214655.1	GI:7288502	
KEYWORDS			
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 9257)		
TITLE	Dominio,S.E. and Lowe,J.B.		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 9257)		
AUTHORS	Dominio,S.E. and Lowe,J.B.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-DEC-1999) HHMI, University of Michigan, 1150 W. Med. Ctr. Dr., Ann Arbor, MI 48109-0650, USA		
FEATURES	Location/Qualifiers		
source	1..9257		
	/organism="Mus musculus"		
	/strain="129/Ola"		
	/db_xref="taxon:10090"		
	/chromosome="7"		
	/map="23.2: between Klk1 and Ldh3"		
	/cell_line="E14"		
	/cell_type="ES"		
	<5578..>6716		
mRNA	/gene="Fut1"		
	/product="alpha(1,2)fucosyltransferase FUT1"		
gene	<5578..>6716		
	/gene="Fut1"		
CDS	5583..6716		
	/gene="Fut1"		
	/EC_number="2.4.1.69"		
	/note="glycosyltransferase; similar to the Mus musculus strain ICR sequences deposited at GenBank Accession Numbers Y09863 and AF13533; similar to the Mus musculus strain NIH Swiss sequence deposited at GenBank Accession Number U90553; similar to the Homo sapiens H blood group alpha(1,2)fucosyltransferase FUT1 sequences deposited at GenBank Accession Numbers M35531 and Z69587"		
	/codon_start=1		
	/product="alpha(1,2)fucosyltransferase FUT1"		
	/protein_id="AA045145.1"		
	/db_xref="GI-7288503"		
	/translation="MWTPSRRQICLFLVLCVLSAGSFFHLNGNFFRNLPLSLVICADYHLKSPVAMYLQPHLPLOTSGSPSCPEQSSLSGTWTITPGFRGNMGQYATLLAQLNGRQAQIPEHMAALAPYRISLPALVDPEVDSLTPWQHLVLDHDMSEYFSLHDPYKLKSGFPSCFTFHLLREQLRRREFTLHNLHREGAQLLSGLRLTGPGAGIRPHTFVHVRGDYLRVMNRWKGVGVDRLQQAMDWFARHKQPIFVVTSGMKWCLENTIDSHGDVFAAGNGQSGTGCKDFALLITOCNHTINTITGTFWNAYLAGDTVYLANFTLPDSEFLKIFRPAALPEPWWGVGINADLSPLQAQDFDPWKPDSLFRLV"		
polya_signal	8074..8079		
BASE COUNT	2267 a 2195 c 2580 g 2215 t		
ORIGIN			

Query Match 18.3%; Score 217.2; DB 94; Length 9257;
Best Local Similarity 74.2%; Pred. No. 1.2e-25;
Matches 307; Conservative 0; Mismatches 93; Indels 14; Gaps 2;


```

QY 720 tctgagcagtggtggaacctcccccagggccgaagcccccagctacgtcttcctga----- 771
Db 136767 CCTGAGCATGTAGAGCGCTTTACAGGCGCGAAGCCCGAGCTATGCGTCCTGACTTCCTT 136826

QY 772 -agccagagcgtcttactatgacatcctcttattattattattattattattatttattt 830
Db 136827 GAATCTAGGCGCTGTTCTTTTGGGTTCACCTATTATTATACGGGTATTATCTTATTT 136886

QY 831 attt-----ttttatttttcttactgagataataaagaggtccagagagagat 879
Db 136887 ATTTATTATTAGTTTCTTTTCTTACTTGGTAATAATRAAGAGTCTGAAGAGAAAT 136941

RESULT 12
AK025772
LOCUS Homo sapiens CDNA: FLJ22119 fis, clone HEP18852. 29-SEP-2000
DEFINITION AK025772
ACCESSION AK025772
VERSION 1 GI:10438390
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens hepatoma cell_line:HepG2 CDNA to mRNA, clone_lib:HEP
clone:HEP18852.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
Unpublished (2000)
TITLE 2 (bases 1 to 2550)
JOURNAL Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
AUTHORS Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
TITLE Submitted (29-AUG-2000) to the DDBJ/EMBL/GenBank databases. Sumio
JOURNAL Sugano, Institute of Medical Science, University of Tokyo,
Laboratory of Genome Structure Analysis, Human Genome Center;
Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@leims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 3'- & 3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
Location/Qualifiers
1..2550
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HepG2"
/cell_type="hepatoma"
/clone="HEP18852"
/clone_lib="HEP"
/note="cloning vector pHE188FL3"
BASE COUNT 836 a 477 c 472 g 765 t
ORIGIN

Query Match 9.6%; Score 113.8; DB 89; Length 2550;
Best Local Similarity 87.9%; Pred No. 3.5e-09;
Matches 124; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1050 ctctgcactccagccagcagcgagattccatctcaaaaaataaataataaa 1109
Db 2405 CACTGCACCTCCAGCGCTAGGTGACAGAGTCCCACTCAATAAAAAA 2464

QY 1110 taaataataataataataaaaaaataaaaaaataaaaaaataaaaaa 1169
Db 2465 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2524

```

```

QY 1170 aaaaaaaaaaaaaaaaaa 1190
Db 2525 AAAAAAAAAAAAAAAAAA 2545

```

```

RESULT 13
AC015686
LOCUS

```

```

DEFINITION AC015686 167403 bp DNA HTG 26-MAY-2000
Homo sapiens chromosome 11 clone RP11-15G21 map 11, WORKING DRAFT
SEQUENCE, 14 unordered pieces.
AC015686

```

```

VERSION AC015686.3 GI:8096826
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.

```

```

ORGANISM

```

```

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE 1 (bases 1 to 167403)

```

```

AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.

```

```

TITLE Homo sapiens chromosome 11, clone RP11-15G21

```

```

JOURNAL Unpublished

```

```

REFERENCE 2 (bases 1 to 167403)

```

```

AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collamore,A.,
Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Hearford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye.W.J., Zimmer,A. and Zody,M.
Direct Submission

```

```

TITLE

```

```

JOURNAL

```

```

COMMENT

```

```

Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 26, 2000 this sequence version replaced gi:6751811.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center

```

```

Center: Whitehead Institute/ MIT Center for Genome Research

```

```

Center code: WIBR

```

```

Web site: http://www-seq.wi.mit.edu

```

```

Contact: sequence_submissions@genome.wi.mit.edu

```

```

----- Project Information

```

```

Center project name: L1336

```

```

Center clone name: 15.G.21

```

```

----- Summary Statistics

```

```

Sequencing vector: M13; M77815; 100% of reads

```

```

Chemistry: Dye-terminator Big Dye; 100% of reads

```

```

Assembly program: Phrap; version 0.960731

```

```

Consensus quality: 158650 bases at least Q40

```

```

Consensus quality: 162167 bases at least Q30

```

```

Consensus quality: 164010 bases at least Q20

```

```

Insert size: 170000; agarose-fp

```

```

Insert size: 166103; sum-of-ctgigs

```

```

Quality coverage: 4.7 in Q20 bases; agarose-fp

```

```

Quality coverage: 4.8 in Q20 bases; sum-of-ctgigs

```

```

-----

```

```

* NOTE: This is a 'working draft' sequence. It currently

```

```

* consists of 14 contigs. The true order of the pieces

```

```

* is not known and their order in this sequence record is

```

```

* arbitrary. Gaps between the contigs are represented as

```

```

* runs of N, but the exact sizes of the gaps are unknown.

```

```

* This record will be updated with the finished sequence

```

```

* as soon as it is available and the accession number will

```

```

* be preserved.

```

```

*

```

```

*

```

```

*

```

```

*

```

```

*

```

```

*

```

```

*

```

```

*

```

```

*

```

```

*

```

```

*

```

* 2290 4990: contig of 2701 bp in length
* 4991 5090: gap of 100 bp
* 5091 8713: contig of 3623 bp in length
* 8714 8813: gap of 100 bp
* 8814 12644: contig of 3831 bp in length
* 12645 12744: gap of 100 bp
* 12745 19421: contig of 6677 bp in length
* 19422 19521: gap of 100 bp
* 19522 29886: contig of 10365 bp in length
* 29887 29986: gap of 100 bp
* 29987 40001: contig of 10015 bp in length
* 40002 40101: gap of 100 bp
* 40102 55470: contig of 15369 bp in length
* 55471 55570: gap of 100 bp
* 55571 70654: contig of 15084 bp in length
* 70655 70754: gap of 100 bp
* 70755 84201: contig of 13447 bp in length
* 84202 84301: gap of 100 bp
* 84302 99373: contig of 15072 bp in length
* 99374 99473: gap of 100 bp
* 99474 115229: contig of 15756 bp in length
* 115230 115329: gap of 100 bp
* 115330 139575: contig of 24246 bp in length
* 139576 139675: gap of 100 bp
* 139676 167403: contig of 27728 bp in length.

FEATURES

source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone_lib="RPC1-11 Human Male BAC"
1. .2189
/note="assembly_fragment"
2290. .4990
/note="assembly_fragment"
5091. .8713
/note="assembly_fragment"
8814. .12644
/note="assembly_fragment"
12745. .19421
/note="assembly_fragment"
19522. .29886
/note="assembly_fragment"
29887. .40001
/note="assembly_fragment"
clone_end:SP6
vector_side:left
40102. .55470
/note="assembly_fragment"
55571. .70654
/note="assembly_fragment"
clone_end:T7
vector_side:left
70755. .84201
/note="assembly_fragment"
84302. .99373
/note="assembly_fragment"
99474. .115229
/note="assembly_fragment"
115330. .139575
/note="assembly_fragment"
139676. .167403
/note="assembly_fragment"
BASE COUNT 49771 a 33781 c 35092 g 47455 t 1304 others
ORIGIN

Query Match 9.2%; Score 109.6; DB 63; Length 167403;
Best Local Similarity 85.8; Pred. No. 1.5e-08;
Matches 121; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Oy 1050 ctctgcactccagccagccagcagcgagattccatctcaaaaaataataataa 1109
Db 19277 CACTGCACCTCCAGCCTGGCAACAACGACGACTCTCTCTCANAAAAA 19336
Oy 1110 taataataataataataataataataataataataataataataataa 1169
Db 19337 AAAAAA 19396
Oy 1170 aaaaaaaaaaaaaaaaaa 1190
Db 19397 AAGAAAAA 19417

RESULT 14

AF113689 AF113689 1298 bp mRNA PRI 03-FEB-2000
LOCUS Homo sapiens clone FLB4701.
DEFINITION
ACCESSION AF113689
VERSION AF113689.1 GI:6855618
KEYWORDS FLI_CDNA.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE
JOURNAL 1 (bases 1 to 1298)
REFERENCE 2 (bases 1 to 1298)
AUTHORS Zhang,C., Yu,Y., Zhang,S., Ouyang,S., Luo,L., Wei,H., Zhou,G.,
Zhou,W., Bi,J., Zhang,Y., Liu,M. and He,F.

TITLE
JOURNAL Functional prediction of the coding sequences of 32 new genes
deduced by analysis of cDNA clones from human fetal liver
Unpublished

TITLE
JOURNAL Direct Submission
Submitted (14-DEC-1998) Department of Experimental Hematology,
Institute of Radiation Medicine, Beijing Taiping Road 27, Beijing
100850, P. R. China

FEATURES
source
1. .1298
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="FLB4701"
/tissue_type="liver"
/dev_stage="fetus"
/note="HQ1219"

repeat_region 919. .1194
/rpt_family="Alu"

BASE COUNT 484 a 253 c 220 g 341 t
ORIGIN

Query Match 8.6; Score 102.6; DB 89; Length 1298;
Best Local Similarity 80.5; Pred. No. 2.2e-07;
Matches 120; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Oy 1042 gagctcgtctgcactccagccagccagcagcgagattccatctcaaaaaataa 1101
Db 1132 GATTCACCACTGCATCCAGCCTGGCGACAGACTTCGTCTCAAAAAA 1191

Oy 1102 taataataataataataataataataataataataataataataataa 1161
Db 1192 AAAAAA 1251

Oy 1162 aaaaaaaaaaaaaaaaaa 1190
Db 1252 AAAAAA 1280

RESULT 15

AF130087 AF130087 2336 bp mRNA PRI 01-DEC-2000
LOCUS Homo sapiens clone FLB8945 PRO2411 mRNA, complete cds.
DEFINITION
ACCESSION AF130087



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2001, 20:54:44 ; Search time 114.92 Seconds
(without alignments)
240.617 Million cell updates/sec

Title: US-09-391-861-2
Perfect score: 1116
Sequence: 1 MDSDETFEHSGLWVSLAG.....SSDPLSMVGFSGRSPSYAS 209

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_16.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_unclassified.*
13: sp_vertebrate.*
14: sp_virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1106	99.1	209	4 Q9NSA1	Q9nsal1 mus sapien
2	803.5	72.0	210	11 Q9JUN1	Q9jnl1 mus musculu
3	251.5	22.5	251	11 Q9EPC2	Q9epc2 mus musculu
4	228.5	20.5	251	4 Q9GZV9	Q9gzv9 homo sapien
5	218.5	19.6	227	13 Q9DDN0	Q9ddn0 gallus gall
6	171	15.3	206	13 Q9YGD8	Q9ygd8 oncorhynchu
7	168	15.1	196	13 Q9YH31	Q9yh31 notophthalm
8	165	14.8	212	11 Q9ESL9	Q9esl9 mus musculu
9	159.5	14.3	191	13 Q9DFC9	Q9dfc9 brachydanio
10	158	14.2	211	4 Q9NP95	Q9np95 homo sapien
11	137	14.1	212	11 Q9EST9	Q9est9 rattus norv
12	135	13.9	237	13 Q9IAI6	Q9iai6 gallus gall
13	153	13.7	245	13 Q9W6A2	Q9w6a2 gallus gall
14	150	13.4	208	13 Q9PVY1	Q9pvy1 xenopus lae
15	149	13.4	252	11 Q9G096	Q9g096 rattus norv
16	147	13.2	253	13 Q9IAI5	Q9iai5 gallus gall
17	146	13.1	192	11 Q9ERW3	Q9erw3 rattus norv
18	145	13.0	199	13 Q9IAI3	Q9iai3 gallus gall
19	144.5	12.9	207	11 Q9ERQ5	Q9erq5 mus musculu

20 144 12.9 212 13 042407
21 143.5 12.9 213 6 Q9NIB9
22 141 12.6 163 11 Q9JHL9
23 141 12.6 170 4 Q9HCT0
24 140.5 12.6 207 11 Q9ESL8
25 140 12.5 194 6 F79150
26 135 12.1 185 11 Q9ERN5
27 135 12.1 192 4 Q95830
28 134 12.0 194 6 Q9N198
29 131 11.7 162 11 Q9ESS2
30 131 11.7 243 13 Q9W6A1
31 130.5 11.7 181 13 Q9IAI7
32 128 11.5 106 6 Q9NIS8
33 125 11.2 146 13 Q07659
34 124.5 11.2 425 5 Q076831
35 123.5 11.1 770 5 P91672
36 123.5 11.1 770 5 Q9VDT9
37 123 11.0 101 13 P79706
38 123 11.0 127 4 Q99517
39 121 10.8 130 6 Q07767
40 121 10.8 196 4 P78443
41 118.5 10.6 1469 4 Q9HAR3
42 118.5 10.6 1474 4 Q94910
43 118 10.6 440 2 Q9EX20
44 117.5 10.5 1515 11 Q88917
45 117 10.5 108 6 Q9NIS7

ALIGNMENTS

RESULT 1
Q9NSA1
ID Q9NSA1 PRELIMINARY; PRT: 209 AA.
AC Q9NSA1;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE FGF-21.
GN FGF-21.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20461777; PubMed=10858549;
RA Nishimura T., Nakatake Y., Konishi M., Itoh N.;
RT "Identification of a novel FGF, FGF-21, preferentially expressed in the liver."
RL Biochim. Biophys. Acta 1492:203-206(2000).
DR EMBL; AB021975; BAA99415.1; -.
DR InterPro; IPR002209; -.
DR InterPro; IPR002348; -.
DR Pfam; PF00167; FGF; 1.
DR PRINTS; PR00263; HBGFGF.
DR PRINTS; PR00262; IL1HBGF.
DR SMART; SM00442; FGF; 1.
SQ SEQUENCE 209 AA; 27925C52A0023823 CRC64;

Query Match 99.1%; Score 1106; DB 4; Length 209;
Best Local Similarity 99.5%; Pred. No. 2.9e-89;
Matches 208; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDSDETFEHSGLWVSLAGLLGACQAHIPDPSPLQFGQVQRQRYLYTDDAQQTEAH 60
Db 1 MDSDETFEHSGLWVSLAGLLGACQAHIPDPSPLQFGQVQRQRYLYTDDAQQTEAH 60
QY 61 LEIRDGTGVGAADQSPESILQLKALKPGVITQILGVKTSRFLCQRPDGCALYGSLLHFDEA 120
Db 61 LEIRDGTGVGAADQSPESILQLKALKPGVITQILGVKTSRFLCQRPDGCALYGSLLHFDEA 120

```

QY 121 CSFRELLEDGYNVYQSEAHGLPLHLPGNKSHPHDPAPRGPARFPLPGLPPAPPEPGI 180
Db 121 CSFRELLEDGYNVYQSEAHGLPLHLPGNKSHPHDPAPRGPARFPLPGLPPAPPEPGI 180
QY 181 LAPQPDVGVSSDPLSMVGPQGRSPSYAS 209
Db 181 LAPQPDVGVSSDPLSMVGPQGRSPSYAS 209

RESULT 2
Q9JUN1 PRELIMINARY; PRT; 210 AA.
AC Q9JUN1;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE FGF-21.
GN FGF-21.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20461777; PubMed=10858549;
RA Nishimura T., Nakatake Y., Konishi M., Itoh N.;
RT "Identification of a novel FGF, FGF-21, preferentially expressed in
the liver.";
RL Blochim. Biophys. Acta 1492:203-206(2000).
DR EMBL; AB025718; BA99416.1; -.
DR InterPro; IPR002209; -.
DR InterPro; IPR002348; -.
DR Pfam; PF00167; FGF; 1.
DR PRINTS; PR00263; HBGFFGF.
DR PRINTS; PR00262; ILIHGFG.
DR SMART; SM00442; FGF; 1.
SQ SEQUENCE 210 AA; 23237 MW; AE02AABA6477E6F0 CRC64;

Query Match 72.0%; Score 803.5; DB 11; Length 210;
Best Local Similarity 79.4%; Pred. No. 8.3e-63;
Matches 158; Conservative 10; Mismatches 30; Indels 1; Gaps 1;

QY 12 GLWVS-VIAGLLGACQAHPIPDSSPLLQFGGQVRQRYLYTDDAQOTEAHLEIREDTG 70
Db 12 GLWVRLLLAVELLYQYQAYPIPDSSPLLQFGGQVRQRYLYTDDAQOTEAHLEIREDTG 71
QY 71 GAADSPESLLQLKALKPGVIOILGVKTSRFLCQRPDQALYGSLLHFDPEACSFRELLED 130
Db 72 GAHRSPESELLEKALKPGVIOILGVKASRFLCQRPDQALYGSLLHFDPEACSFRELLED 131
QY 131 GYNVYQSEAHGLPLHLPGNKSHPHDPAPRGPARFPLPGLPPAPPEPGIILAPQPDVGS 190
Db 132 GYNVYQSEAHGLPLHLPGNKSHPHDPAPRGPARFPLPGLPPAPPEPGIILAPQPDVGS 191
QY 191 SDPLSMVGPQGRSPSYAS 209
Db 192 SDPLSMVGPQGRSPSYAS 210

RESULT 3
Q9EPC2 PRELIMINARY; PRT; 251 AA.
AC Q9EPC2;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE FGF-23 (FGF23).
GN FGF23.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11032749;
RA Yamashita T., Yoshioka M., Itoh N.;
RT "Identification of a novel fibroblast growth factor, FGF-23,
preferentially expressed in the ventrolateral thalamic nucleus of the
brain.";
RL Biochem. Biophys. Res. Commun. 277:494-498(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=WHOLE EMBRYO;
RX MEDLINE=20517346; PubMed=11062477;
RA White K.E., Evans W.E., O'Riordan J.L.H., Speer M.C., Econs M.J.,
RA Lorenz-Deptieux B., Grabowski M., Meitinger T., Strom T.M.;
RT "Autosomal dominant hypophosphataemic rickets is associated with
mutations in FGF23.";
RL Nat. Genet. 26:345-348(2000).
DR EMBL; AB037889; BAB13478.1; -.
DR EMBL; AF263536; AAG09916.1; -.
SQ SEQUENCE 251 AA; 27757 MW; 110C1F2C735DC360 CRC64;

Query Match 22.5%; Score 251.5; DB 11; Length 251;
Best Local Similarity 32.1%; Pred. No. 1.9e-14;
Matches 76; Conservative 23; Mismatches 85; Indels 53; Gaps 9;

QY 12 GLWVS-VIAGLLGACQ---AHPIPDSSPLLQFGGQVRQRYLYTDDAQOTEAHLEIREDT 66
Db 3 GTCRLRLVGLCTVCSIGCTARAYPDTPSLLGSNWSLT---HLYTATA-RTSYHLQIHRD 58
QY 67 GTVGAADSPESLLQLKALKPGVIOILGVKTSRFLCQRPDQALYGSLLHFDPEACSFREL 126
Db 59 GHVDGTPHOTIYSALMITSEDAGSVVITGAMTRRFLCMLHGNIFGSLHFSPEACKFROW 118
QY 127 LLEDGYNVYQSEAHGLPLHL-----PGNKS----- 152
Db 119 TLENGDYVLSQKHLYLVSLGRAKRIQFQTPNPFPSPQFLARRNEVPLLHFTVVRPRHT 178
QY 153 --HRDPAPRGPARFL-PLPGLPPAP---PEPPGILAPQPDVGVSSDPLSMVGPQSG 202
Db 179 RSAEDIPERDPLNLVLPKPRAPFPVPVSCSRELPSAEEGFP---AASDPLGLVLRGRG 232

RESULT 4
Q9GV9 PRELIMINARY; PRT; 251 AA.
AC Q9GV9;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE FGF-23 (FGF23).
GN FGF23.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11032749;
RA Yamashita T., Yoshioka M., Itoh N.;
RT "Identification of a novel fibroblast growth factor, FGF-23,
preferentially expressed in the ventrolateral thalamic nucleus of the
brain.";
RL Biochem. Biophys. Res. Commun. 277:494-498(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20517346; PubMed=11062477;
RA White K.E., Evans W.E., O'Riordan J.L.H., Speer M.C., Econs M.J.,
RA Lorenz-Deptieux B., Grabowski M., Meitinger T., Strom T.M.;
RT "Autosomal dominant hypophosphataemic rickets is associated with
mutations in FGF23.";
RL Nat. Genet. 26:345-348(2000).
DR EMBL; AB037973; BAB13477.1; -.

```

RT "Putative Newt Fibroblast Growth Factor-4";
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U76998; AAC98812.1; -
 DR HSSP; P09038; 1BFF.

```
DR InterPro; IPR001064; -.
DR InterPro; IPR002209; -.
DR InterPro; IPR002348; -.
DR Pfam; PF00167; FGF; 1
DR PRINTS; PR00263; HBGFFGF.
DR PRODOM; PD000831; -.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
DR PROSITE; PS00247; HBG_FGF; 1.
DR SMART; SM00442; FGF; 1.
DR SEQUENCE 196 AA; 22033 MW; AC4688CD989C6EAF CRC64;

Query Match 15.1%; Score 168; DB 13; Length 196;
Best Local Similarity 28.9%; Pred. No. 2.8e-07;
Matches 44; Conservative 27; Mismatches 45; Indels 36; Gaps 5;

QY 17 VLAGLLGACQAHPIPD--SSPLLQFGGQVRQRYLYTDD-----AQTEA----- 59
DB 15 VLLGLLSGLARCLPMPSPGNGTLEWSWET-----LYSQLARLAGGQRTDAHRYGEYLLGI 70
QY 60 -----HLEIREDTGVTGAADQSPESILQLKALKPGVQIILGVKTSRFLCQR 105
DB 71 KLRRLRYCNVGIQFLQVLPDGRIGHMHSERYSLLEISPVGVCMFGVQSGLEFLAMN 130
QY 106 PDGALYGSILHFDPEACSFRELLLEDGYNYVQS 137
DB 131 SKGRLEFGSKYFSDE-CKFKEMLLPNYAYES 161

RESULT 8
Q9ESL9 PRELIMINARY; PRT; 212 AA.
AC Q9ESL9;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE FIBROBLAST GROWTH FACTOR 20.
GN FGF20.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Itoh N.;
RT "Mus musculus mRNA for FGF-20 (FGF20).";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB049218; BAB16406.1; -.
SQ SEQUENCE 212 AA; 23659 MW; 174DBCE8915B69EF CRC64;

Query Match 14.8%; Score 165; DB 11; Length 212;
Best Local Similarity 33.5%; Pred. No. 5.7e-07;
Matches 54; Conservative 18; Mismatches 63; Indels 26; Gaps 6;

QY 32 PDSPLLQFGGQVRQRYLYTDDAQOTEAHLEIREDTGVTGA-ADQSPESILQLKALKPGV 90
DB 52 PGCSVELAHLGLIRRLQYLC----RTGFLHLQILPDGTQGTQDHSFLGILEFISVAVGL 107
QY 91 IOILGVKTSRFLCQRPDGALYGSILHFDPEACSFRELLLEDGYNYVQSPAHLPLHLPGNK 150
DB 108 VSRGVDSGLVGMNDKGLYSEKLTSE-CIFRQEQFEENYNTYSSN-----IYKHGNT 161
QY 151 -----SPHRPAPRGPAR-----FLPLGLPPAPPE 176
DB 162 GRRYFVNLKDGTPRDGARSRRKROKFTFLRPVDPERVPE 202

RESULT 9
Q9DFC9 PRELIMINARY; PRT; 191 AA.
AC Q9DFC9;
```

```
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE FIBROBLAST GROWTH FACTOR 4 PRECURSOR.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Rasbora; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Draper B.W.; Kimmel C.B.;
RT "FGF4 Expression is Associated with Numerous Developmentally Important Signaling Centers During Zebrafish Embryogenesis.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF283555; AAG13950.1; -.
SQ SEQUENCE 191 AA; 21203 MW; D2B5F2FAD44B2289 CRC64;

Query Match 14.3%; Score 159.5; DB 13; Length 191;
Best Local Similarity 26.3%; Pred. No. 1.5e-06;
Matches 51; Conservative 27; Mismatches 69; Indels 47; Gaps 6;

QY 11 SGLWVSVLAGLLGACQAHPIPD-----SSPLLQFGGQVRQ----- 46
DB 5 SALLPILVLGLMTSSVRCAPLPGGHSVPVRRWETLYSRSLARIPGCKRDISRDSYLTG 64
QY 47 ----RYLYTDDAQOTEAHLEIREDTGVTGAADQSPESILQLKALKPGVQIILGVKTSRFL 102
DB 65 IKLRRLRYCNV--IGFHLQVLPGGKITGVHNRYSLLLEISPVGVVTLFGVRSGLFV 122
QY 103 QCRPDGALYGSILHFDPEACSFRELLLEDGYNYVQSPAHLPLHLPG-----NKSP 152
DB 123 AMNSKGLYSGEQFTNE-CKFREKLLANNYAYESLAH-----PGMYIGLSKAGTKKG 175
QY 153 HRDPAPRGPARFLP 166
DB 176 NRVSTSTMTWTFLP 189

RESULT 10
Q9NP95 PRELIMINARY; PRT; 211 AA.
AC Q9NP95;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE FGF-20 (FIBROBLAST GROWTH FACTOR 20) (FGF-20).
GN FGF-20.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20374469; PubMed-10913340;
RA Kirikoshi H., Sagara N., Saitoh T., Tanaka K., Sekihara H.,
RA Shikawa K., Katoh M.;
RT "Molecular cloning and characterization of human FGF-20 on chromosome 8p21.3-p22.2.";
RL Biochem. Biophys. Res. Commun. 274:337-343(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Itoh N.;
RT "Human FGF-20.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB044277; BAB03633.1; -.
DR EMBL; AB030648; BAB03530.1; -.
DR InterPro; IPR002209; -.
DR InterPro; IPR002348; -.
DR Pfam; PF00167; FGF; 1.
DR PRINTS; PR00263; HBGFFGF.
DR PRINTS; PR00262; ILLHBGF.
```



```
DR PROSITE: PS00247; HBGF_FGF; 1.
DR SMART: SM00442; FGF; 1
SQ SEQUENCE 211 AA; 23498 MW; AB04608C16060CC1 CRC64;

Query Match 14.2%; Score 158; DB 4; Length 211;
Best Local Similarity 33.1%; Pred. No. 2.3e-06;
Matches 53; Conservative 18; Mismatches 65; Indels 24; Gaps 6;

QY 32 PDSSPLLQFGGQVQRVLYTDDAQTEAHLEIREDTGTVGGA-ADQSPESLLQLKALKPGV 90
DB 52 PGSAQLAHLGLRRRLQYLC-----RTGFHLQLLPDGSVQGRQDHSFGILFEISVAVGL 107
QY 91 IQLGVKTSRFLCQPDGALYGLHFDPEACSFRELLLEDGNYVQSEAH-----GLPLH 145
DB 108 VSRGVDSGLYLGMDKGLYSEKLTSE-CIFRQGFENWNTYSSNIYKHGDTGRRYF 166
QY 146 LPGNKSPHRDPAPRGPAR-----FLPLGLPPAPPE 176
DB 167 VALNK-----DGTPRDGARSKRHQKTFHLP RPVDPDPERVPE 202

RESULT 11
Q9EST9 PRELIMINARY; PRT; 212 AA.
AC Q9EST9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE FGF-20.
GN FGF-20.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC PubMed=11032730;
RX Tissue=ADULT BRAIN;
RA Ohmachi S., Watanabe Y., Mikami T., Kusu N., Ibi T., Akaike A.,
RA Itoh N.;
RA "FGF-20, a novel neurotrophic factor, preferentially expressed in the
RT substantia nigra pars compacta of rat brain.";
RL Biochem. Biophys. Res. Commun. 277:355-360(2000).
DR EMBL: AB020021; BAB13763.1; -.
SQ SEQUENCE 212 AA; 23537 MW; 4F858BEFE772B977 CRC64;

Query Match 14.1%; Score 157; DB 11; Length 212;
Best Local Similarity 33.8%; Pred. No. 2.8e-06;
Matches 54; Conservative 16; Mismatches 66; Indels 24; Gaps 6;

QY 32 PDSSPLLQFGGQVQRVLYTDDAQTEAHLEIREDTGTVGGA-ADQSPESLLQLKALKPGV 90
DB 52 PGSEVELAHLGLRRRLQYLC-----RTGFHLQLLPDGSVQGRQDHSFGILFEISVAVGL 107
QY 91 IQLGVKTSRFLCQPDGALYGLHFDPEACSFRELLLEDGNYVQSEAH-----GLPLH 145
DB 108 VSRGVDSGLYLGMDKGLYSEKLTSE-CIFRQGFENWNTYSSNIYKHGDTGRRYF 166
QY 146 LPGNKSPHRDPAPRGPAR-----FLPLGLPPAPPE 176
DB 167 VALNK-----DGTPRDGARSKRHQKTFHLP RPVDPDPERVPE 202

RESULT 12
Q9IAI6 PRELIMINARY; PRT; 237 AA.
AC Q9IAI6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE FIBROBLAST GROWTH FACTOR HOMOLOGOUS FACTOR 4 ISOFORM 1A (FRAGMENT).

GN FHF-4.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=9847253;
RA Munoz-Sanjuan I., Simandl B.K., Fallon J.F., Nathans J.;
RT "Expression of chicken fibroblast growth factor homologous factor
(PHF)-1 and of differentially spliced isoforms of PHF-2 during
RT development and involvement of PHF-2 in chicken limb development.";
RL Development 126:409-421(1999).
RN [2]
RP SEQUENCE FROM N.A.
```

RC	TISSUE=BRAIN;
RX	MEDLINE=201112823; PubMed=10644718;
RA	Munoz-Sanjuan I., Smallwood P.M., Nathans J.;
RT	"Isoform Diversity among Fibroblast Growth Factor Homologous Factors
IS	Is Generated by Alternative Promoter Usage and Differential
RL	Splicing.";
RL	J. Biol. Chem. 275:2589-2597(2000).
DR	EMBL; AF108755; AAD21576.1; -.
DR	HSP; AF199611; AAF31398.1; -.
DR	HSSP; P03968; IBAR.
DR	InterPro; IPR002209; -.
DR	InterPro; IPR002348; -.
DR	Pfam; PF00167; FGF; 1.
DR	PRINTS; PR00263; HBGFFGF.
DR	PRINTS; PR00262; ILIHGFG.
DR	PROSITE; PS00247; HBGF_FGF; 1.
DR	SMART; SM00442; FGF; 1.
SQ	SEQUENCE 245 AA; 27605 MW; DA0DAFEDBA1B8F28 CRC64;

Query Match	13.7%; Score 153; DB 13; Length 245;
Best Local Similarity	29.8%; Pred. No. 7.5e-06;
Matches 50; Conservative 23; Mismatches 67; Indels 28; Gaps 6;	

QY	24	GACQAHPIDPSLLQFGGVQRVY-----LYTDDAQTEAHLEIREDTGVGGAADQ	75
DB	36	GSCDKNLNVSRVKLFSGKRRRPPQLKGIVTKLYSGVHLQLQAOGTIDGTKEE	95
QY	76	-SPESLLQALKPGVIOILGVKTSRFQCQRPDGALYGLSHFDPEACSFRELLLEDGYNV	134
DB	96	DSSYTFLNLIPIVGLRVVAIQGVTKLYLANSEGGYLTVEHFTPE-CFKKESIFENYYVT	154
QY	135	Y-----QSEAHGLPLHL-----PGNKSPHRDPAPRGPARPLP	168
DB	155	YSSMIYRQQQSGRWYLGNKEGEIMGNHVKKNP-----AHLPKP	198

RESULT 14	
Q9PVY1	
ID	Q9PVY1 PRELIMINARY; PRT; 208 AA.
AC	Q9PVY1;
DT	01-MAY-2000 (TrEMBLrel. 13, Created)
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT	01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE	XFGF-20.
GN	XFGF-20.
OS	xenopus laevis (African clawed frog).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Pipidae;
OC	Xenopodinae; Xenopus.
OX	NCBI_TaxID=8355;
RN	[1]
RX	SEQUENCE FROM N.A.
RP	MEDLINE=99373151; PubMed=10441498;
RA	Koga C., Adati N., Nakata K., Mikoshiba K., Furuhata Y., Sato S.,
RA	Tei H., Sakaki Y., Kurokawa T., Shikokawa K., Yokoyama K.K.;
RT	"Characterisation of a novel member of the FGF family, xFGF-20, in
RT	Xenopus laevis.";
RL	Biochem. Biophys. Res. Commun. 261:756-765(1999).
DR	EMBL; AB012615; BAA83474.1; -.
DR	HSP; P05230; 2AFG.
DR	InterPro; IPR002209; -.
DR	InterPro; IPR002348; -.
DR	Pfam; PF00167; FGF; 1.
DR	PRINTS; PR00263; HBGFFGF.
DR	PRINTS; PR00262; ILIHGFG.
DR	PROSITE; PS00247; HBGF_FGF; 1.
DR	SMART; SM00442; FGF; 1.
SQ	SEQUENCE 208 AA; 23438 MW; 268881D36E757D4D CRC64;

Query Match	13.4%; Score 150; DB 13; Length 208;
Best Local Similarity	31.5%; Pred. No. 1.1e-05;

Result No.	Query	Score	Match	Length	DB	ID	Description
1	273	24.5	216	1	FGF1_HUMAN	O95750	homo sapien
2	206.5	18.5	218	1	FGF1_MOUSE	O35622	mus musculus
3	179	16.0	245	1	FGF3_MOUSE	P05524	mus musculus
4	177	15.9	206	1	FGF4_BOVIN	P48003	bos taurus
5	176.5	15.8	208	1	FGF6_HUMAN	P10767	homo sapien
6	174	15.6	187	1	FGF4_XENLA	P48805	xenopus lae
7	170.5	15.3	208	1	FGF6_MOUSE	P21658	mus musculus
8	170.5	15.3	239	1	FGF3_HUMAN	P11487	homo sapien
9	168	15.1	266	1	FGF5_RAT	P48807	rattus norv
10	165	14.8	194	1	FGF4_CHICK	P48804	gallus gall
11	163.5	14.7	192	1	FGF4_XENLA	P48806	xenopus lae
12	163.5	14.7	206	1	FGF4_HUMAN	P08620	homo sapien
13	161	14.4	264	1	FGF5_MOUSE	P15656	mus musculus
14	160.5	14.4	256	1	FGF3_BRARE	P15656	mus musculus
15	155.5	13.9	155	1	FGF1_MESAU	P48802	brachydanio
16	155	13.9	247	1	FGF6_HUMAN	P34004	mesocricetu
17	154.5	13.8	155	1	FGF1_MOUSE	Q92915	homo sapien
18	154	13.8	208	1	FGF9_HUMAN	P10935	mus musculus
19	154	13.8	208	1	FGF9_RAT	P31371	homo sapien
20	153.5	13.8	202	1	FGF4_MOUSE	P36364	rattus norv
21	153.5	13.8	208	1	FGF9_MOUSE	P11403	mus musculus
22	153	13.7	268	1	FGF5_HUMAN	P54130	mus musculus
23	153	13.4	208	1	FGF4_HUMAN	P12034	homo sapien
24	150	13.4	215	1	FGF4_RAT	O15520	homo sapien
25	150	13.4	245	1	FGF6_MOUSE	P70492	rattus norv
26	149	13.4	247	1	FGF6_MOUSE	P70377	mus musculus
27	148	13.3	209	1	FGF4_MOUSE	P70377	mus musculus
28	146.5	13.1	155	1	FGF1_HUMAN	O35565	mus musculus
29	146	13.1	237	1	FGF3_XENLA	P05230	homo sapien
30	144.5	12.9	155	1	FGF1_BOVIN	P36386	xenopus lae
31	143.5	12.9	152	1	FGF1_PIG	P30968	bos taurus
32	143.5	12.9	207	1	FGF6_HUMAN	P20002	sus scrofa
33	142	12.7	220	1	FGF3_CHICK	O43320	homo sapien

SEQUENCE FROM N.A.	
Xie M.-H., Holcomb I., Deuel B., Dowd P., Huang A., Vagts A.,	
Foster J., Brush J., Gu Q., Liang J., Hillan K., Goddard A.,	
Gurney A.L.;	
"FGF-19, a novel fibroblast growth factor with unique specificity for	
FGFR4";	
Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.	
-!- FUNCTION: MAY BE INVOLVED IN BRAIN DEVELOPMENT DURING	
EMBRYOGENESIS.	
-!- SUBCELLULAR LOCATION: SECRETED.	
-!- TISSUE SPECIFICITY: EXPRESSED ONLY IN FETAL BRAIN.	
-!- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.	

This SWISS-PROT entry is copyright. It is produced through a collaboration	
between the Swiss Institute of Bioinformatics and the EMBL outstation -	
the European Bioinformatics Institute. There are no restrictions on its	
use by non-profit institutions as long as its content is in no way	
modified and this statement is not removed. Usage by and for commercial	
entities requires a license agreement (See http://www.isb-sib.ch/announcement/	
or send an email to license@isb-sib.ch).	

EMBL; AB018122; BAA75500.1; -;	
EMBL; AF110400; AAD45973.1; -;	
MIM; 603891; -;	
HSP; P09038; 2BFH.	
InterPro; IPR002209; -;	
InterPro; IPR002348; -;	
Pfam; PF00167; FGF; 1.	
PRINTS; PR00262; IL1HBGF.	
PRINTS; PR00263; HBGF.FGF.	
PROSITE; PS00247; HBGF_FGF; 1.	
Growth factor; Signal.	
SIGNAL	1 22
FT CHAIN	23 216
FT	POTENTIAL.
FT	FTFIBROBLAST GROWTH FACTOR-19.

```
SQ SEQUENCE 216 AA; 24002 MW; E0BCBC9C220F9832 CRC64;

Query Match 24.5%; Score 273; DB 1; Length 216;
Best Local Similarity 36.8%; Pred. No. 9.9e-14;
Matches 78; Conservative 27; Mismatches 81; Indels 26; Gaps 9;

QY 11 SGLWSVLVAGLLGACQAHPTDSSPLLOF--GGOVORYLYTDDAQ-QTEAHLEIREDDG 67
   :|||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 14 AGLWLAV-AGRPPLA-----FSDAGPHVHYGWDPIRLRHLYTSGPHGLSCFLIRADG 66
   :|||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 68 TVGGADQSPESLLQKALKPGVIOILGVKTSRFLCQPDGALYGLSHFDPEACSFRELL 127
   :|||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 67 VYDCARGSAHSLLEIKVALETKVAIKGVHVSRYLCMGADKMGQGLLOYSEDCAFEEI 126
   :|||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 128 LEDGYNVYQSEAHGLPLHLPNGKNSPHRDPAPRPFPLP-----LGPLPAPPEPPI 180
   :|||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 127 RPDGYNVYRSEKHLRPLVSLSSAKQ-RQLYKNRG---FLPLSHFLPLMPVPEEPDLRGH 182
   :|||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 181 ----LAPQPPDVGGSDPLSMV-GPSQGRSPSY 207
   :|||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 183 LESDMFSPSLETSDMDPGLVLTGLEAVRSPSF 214
   :|||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 2
FGFF_MOUSE
ID FGF3_MOUSE STANDARD; PRT; 218 AA.
AC O35622;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE FIBROBLAST GROWTH FACTOR-15 PRECURSOR (FGF-15).
GN FGF15.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97454242; PubMed=9310317;
RA McWhirter J.R., Goulding M., Weiner J.A., Chun J., Murre C.;
RT "A novel fibroblast growth factor gene expressed in the developing
RT nervous system is a downstream target of the chimeric homeodomain
RT oncoprotein E2A-Pbx1."
RL Development 124:3221-3232(1997).
CC -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN REGULATING CELL DIVISION
CC AND PATTERNING WITHIN SPECIFIC REGIONS OF THE EMBRYONIC BRAIN,
CC SPINAL CORD AND SENSORY ORGANS.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE DEVELOPING BRAIN.
CC -!- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF007268; AAB63197.1; -.
CC DR MGD; MGI:1096383; Fgf15.
CC DR InterPro; IPR002209; -.
CC DR InterPro; IPR002348; -.
CC DR Pfam; PF00167; FGF; 1.
CC DR PRINTS; PR00262; ILLHGF.
CC DR PROSITE; PS00247; HBGF_FGF; 1.
CC DR Growth factor; Signal.
CC FT CHAIN 1 25 POTENTIAL.
CC FT CHAIN 26 218 FIBROBLAST GROWTH FACTOR-15.
CC SQ SEQUENCE 218 AA; 25236 MW; A96B0D771FE125A5 CRC64;
```

```
Query Match 18.5%; Score 206.5; DB 1; Length 218;
Best Local Similarity 33.8%; Pred. No. 8.1e-09;
Matches 53; Conservative 29; Mismatches 70; Indels 5; Gaps 4;

QY 13 LWSVLVAGLLGACQAHPTDSSPLLOF-GQV-RORYLYTDDAQTEAHLEIREDDGTVG 70
   :|||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 19 LWLAVSGRPL--AQSQSVSEDFLYGKTRQLYLYSAGPYVSNCFLRSDGSVD 76
   :|||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 71 GADQSPESLLQKALKPGVIOILGVKTSRFLCQPDGALYGLSHFDPEACSFRELL 130
   :|||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 77 CEEDQERNLLEFRAVALKTAIAIKDVSSRYLCMDGKIYGLIRYSEEDCTPREMDCL 136
   :|||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 131 GYNVYQSEAHGLPLHLPNGKNSPHRDPAPRPFPLP 167
   :|||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 137 GYNVYRSMKHLHIIFTQAK-PRQLQDQKPSNPIPV 172
   :|||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 3
FGF3_MOUSE
ID FGF3_MOUSE STANDARD; PRT; 245 AA.
AC P05524;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE INT-2 PROTO-ONCOGENE PROTEIN PRECURSOR (FIBROBLAST GROWTH FACTOR-3)
DE (FGF-3) (HBGF-3).
GN FGF3 OR FGF-3 OR INT-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86247582; PubMed=3013624;
RA Moore R., Casey G., Brookes S., Dixon M., Peters G., Dickson C.;
RT "Sequence, topography and protein coding potential of mouse int-2: a
RT putative oncogene activated by mouse mammary tumour virus."
RL EMBO J. 5:919-924(1986).
RN [2]
RP SEQUENCE OF 1-73 FROM N.A.
RX MEDLINE=88296404; PubMed=2841106;
RA Smith R., Peters G., Dickson C.;
RT "Multiple RNAs expressed from the int-2 gene in mouse embryonal
RT carcinoma cell lines encode a protein with homology to fibroblast
RT growth factors."
RL EMBO J. 7:1013-1022(1988).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=91193291; PubMed=1964688;
RA Dickson C., Acland P., Smith R., Dixon M., Deed R., McAllan D.,
RA Walther W., Fuller-Pace F., Kiefer P., Peters G.;
RT "Characterization of int-2: a member of the fibroblast growth factor
RT family."
RL J. Cell Sci. Suppl. 13:87-96(1990).
CC -!- FUNCTION: COULD BE INVOLVED IN EAR DEVELOPMENT.
CC -!- INDUCTION: BY INTEGRATION OF MOUSE MAMMARY TUMOR VIRUS.
CC -!- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y00848; CAA68767.1; -.
CC DR PIR; A23930; TVNST2.
CC DR HSSP; P09038; 2BFH.
CC DR MGD; MGI:95517; Fgf3.
CC DR InterPro; IPR002209; -.
CC DR InterPro; IPR002348; -.
CC DR Pfam; PF00167; FGF; 1.
CC DR
```

```

SQ SEQUENCE 206 AA; 22041 MW; F9B7A86066E56613 CRC64;

Query Match 15.9%; Score 177; DB 1; Length 206;
Best Local Similarity 32.1%; Pred. No. 1,1e-06;
Matches 52; Conservative 23; Mismatches 75; Indels 12; Gaps 5

QY 14 WWSVLGALLGACQAHPIPDSSPLLOFGGV-----RQRYLYTDDAQOQTAHLEIREDTG 68
DB 46 WESIVARSLLAGLPVAAQKEAAVSGAGDVLGKIKRLRRLYCNVG--IGFHLQVLPDGR 103
QY 69 VGGAAQDSPESSLLQALKAPGVQIQLGKTSFELCQRPDGLGALGSLHFDPEACSFRELLL 128
DB 104 IGGVHADTSDSLELSFVGRGVVSGFVASRFFVAMSSRGRLYSGPFTFDE-CRFREILL 162
QY 129 EDGYNVYQSEAH-GLPLHLPGN---KSPHRDPAPRGPARFLP 166
DB 163 PNYNAYECDRHPGWFIALSKNGKAKGNRVSPMKVTHFLP 204

RESULT 5
FGF6_HUMAN
ID FGF6_HUMAN STANDARD; PRT; 208 AA.
AC P10767;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE FIBROBLAST GROWTH FACTOR-6 PRECURSOR (FGF-6) (HSGF-6) (HST-2).
GN FGF6 OR HST2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91360279; PubMed=1866714;
RA Coullier F., Batoz M., Marics I., de Lapeyriere O., Birnbaum D.;
RT "Putative structure of the FGF6 gene product and role of the signal
RT peptide.";
RL Oncogene 6:1437-1444(1991).
RN [2]
RP SEQUENCE OF 11-208 FROM N.A.
RX MEDLINE-92195660; PubMed=1549352;
RA Tida S., Yoshida T., Naito K., Sakamoto H., Katoh O., Hirohashi S.,
RA Sato T., Onda M., Sugimura T., Terada M.;
RT "Human hst-2 (FGF-6) oncogene: cDNA cloning and characterization.";
RL Oncogene 7:303-309(1992).
RN [3]
RP SEQUENCE OF 81-208 FROM N.A.
RX MEDLINE-89201880; PubMed=2649847;
RA Marics I., Adelaide J., Raybaud F., Mattei M.-G., Coullier F.,
RA planche J., de Lapeyriere O., Birnbaum D.;
RT "Characterization of the HST-related FGF.6 gene, a new member of the
RT fibroblast growth factor gene family.";
RL Oncogene 4:335-340(1989).
CC -!- FUNCTION: CAN TRANSFORM NIH 3T3 CELLS. EXHIBITS STRONG MITOGENIC
CC AND ANGIOGENIC PROPERTIES.
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -!- TISSUE SPECIFICITY: LEUKEMIA CELL LINES WITH PLATELET/
CC MEGAKARYOCYTIC DIFFERENTIATION POTENTIAL.
CC -!- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X57075; CAA04358.1; -
CC DR EMBL; X57075; CAA04359.1; ALT INIT.
CC DR EMBL; X57075; CAA04360.1; ALT INIT.

```


use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M92415; AAA62260.1; -
EMBL; M92416; AAA62261.1; -
EMBL; X51552; CAA35925.1; -
HSSP; P09038; 1BLD.
MGD; MGI:95520; Fgf6.
InterPro; IPR002209; -
InterPro; IPR002348; -
Pfam; PF00167; FGF; 1.
PRINTS; PR00263; HBGFFGF.
PRINTS; PR00262; ILLHBGF.
PROSITE; PS00247; HBGF_FGF; 1.
Transforming protein; Oncogene; Growth factor; Mitogen; Signal;
Glycoprotein.
FT SIGNAL 1 ? POTENTIAL.
FT CHAIN ? 208 FIBROBLAST GROWTH FACTOR-6.
FT CARBOHYD 45 45 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFID 90 157 POTENTIAL.
SQ SEQUENCE 208 AA; 22798 MW; BBF7B9301BB79A1B CRC64;

Query Match 15.3%; Score 170.5; DB 1; Length 208;
Best Local Similarity 36.4%; Pred. No. 3.5e-06;
Matches 36; Conservative 21; Mismatches 39; Indels 3; Gaps 2;

QY 42 GQVQRRLYTDDAQTEAHLEIREDTGVGAADQSPESLLQLKALKPGVQILGVKTSRF 101
DB 81 GIKRORRLYCNVVG--IGFHLQVPPDGRISGTHEENPYSLLEISTVGRGVWSLFGVKSALF 138

QY 102 LCPORPDGALYGLSHDPPEACSFRELLLEDGYNNYQSEAH 140
DB 139 IAMNSKGRLLYTPPSHDE-CKEFTLLPNPNYAYESDLY 176

RESULT 8
FGF3_HUMAN STANDARD; PRT; 239 AA.
AC FGF3_HUMAN
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE INT-2 PROTO-ONCOGENE PROTEIN PRECURSOR (FIBROBLAST GROWTH FACTOR-3)
DE (FGF-3) (HBGF-3).
GN FGF3 OR INT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=89239468; PubMed=2470007;
RA Brooks S., Smith R., Casey G., Dickson C., Peters G.;
RT "Sequence organization of the human int-2 gene and its expression in teratocarcinoma cells.";
RL Oncogene 4:429-436(1989).
CC -!- FUNCTION: COULD BE INVOLVED IN EAR DEVELOPMENT.
CC -!- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X14445; CAA32615.1; -
PIR; S04742; S04742.

HSSP; P09038; 2BFH.
MIM; 164950; -
InterPro; IPR002209; -
InterPro; IPR002348; -
Pfam; PF00167; FGF; 1.
PRINTS; PR00262; ILLHBGF.
PRINTS; PR00263; HBGFFGF.
PROSITE; PS00247; HBGF_FGF; 1.
Transforming protein; Oncogene; Growth factor; Mitogen; Signal;
Glycoprotein.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 239 INT-2 PROTO-ONCOGENE PROTEIN.
FT CARBOHYD 65 65 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 239 AA; 26886 MW; 8DBEF17D2B2E3C63 CRC64;

Query Match 15.3%; Score 170.5; DB 1; Length 239;
Best Local Similarity 29.1%; Pred. No. 4e-06;
Matches 50; Conservative 23; Mismatches 58; Indels 41; Gaps 6;

QY 41 GQVQRRLYTDDAQTEAHLEIREDTGVGAADQSPESLLQLKALKPGVQILGVKTSR 100
DB 40 GGAPRRRLKLYC---ATKYHLQLHPSGRVNGSLNSAYSILEITAVEVGIVAIRGLFSGR 95

QY 101 FLCPORPDGALYGLSHDPPEACSFRELLLEDGYNNYQSEAHGLPLHLPKNSPHRDAP-- 158
DB 96 YLAMNKRGLRYASEHSYSAE-CEFERIHELGYNTYASRLYRTVSSSTPGAR---RQPSAER 151

QY 159 -----RGPAR-----FLP-----LPGLPPAPPEPPG 179
DB 152 LMTYSVNGKGRPRGRGFKYRTKQSSLEPLVLDHRDHEMYRQSLGPRPFG 203

RESULT 9
FGF5_RAT STANDARD; PRT; 266 AA.
AC FGF5_RAT
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE FIBROBLAST GROWTH FACTOR-5 PRECURSOR (FGF-5) (HBGF-5).
GN FGF5 OR FGF-5.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar;
RX MEDLINE=96201703; PubMed=8611621;
RA Hattori Y., Yamasaki M., Itoh N.;
RT "The rat FGF-5 mRNA variant generated by alternative splicing encodes a novel truncated form of FGF-5.";
RL Blochim. Biophys. Acta 1306:31-33(1996).
CC -!- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D64085; BAA10966.1; -
HSSP; P09038; 1BFC.
InterPro; IPR002209; -
InterPro; IPR002348; -
Pfam; PF00167; FGF; 1.
PRINTS; PR00262; ILLHBGF.
PRINTS; PR00263; HBGFFGF.
PROSITE; PS00247; HBGF_FGF; 1.
Transforming protein; Oncogene; Growth factor; Mitogen; Signal;
KW


```
SQ SEQUENCE 192 AA; 21903 MW; 2B01B0BB8824E3B3 CRC64;

Query Match
Best Local Similarity 32.4%; Score 163.5; DB 1; Length 192;
Matches 36; Conservative 26; Mismatches 44; Indels 5; Gaps 3;

QY 60 HLEIREDTGVTGGAAQSPESLLQALKPGVIOILGVKTSRFLCQPDGALYGLSHFDPE 119
DQ 81 HIQVLDPGRINGMHNRYSLLETSPEVGVWSLYGKISAMFVAMNAKGLYGRYFN-E 139
QY 120 ACSFRELLEDGYNVQSEAH-GLPLHLPGN---KSPHRDPAPRGPARFLP 166
DQ 140 ECKFEKLLPNPNAYESRYPGMYIALGKNGRTKKNRVSPVTMTLTHFLP 190

RESULT 12
FGF4_HUMAN STANDARD; PRT; 206 AA.
AC P08620;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE FIBROBLAST GROWTH FACTOR-4 PRECURSOR (FGF-4) (HEPARIN SECRETORY
DE TRANSFORMING PROTEIN) (HST-1) (HST) (TRANSFORMING PROTEIN KS3) (HBGF-
DE 4).
GN FGF4 OR HST OR HSTF1 OR KS3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88041096; PubMed=2959959;
RA Yoshida T., Miyagawa K., Odagiri H., Sakamoto H., Little P.F.R.,
RA Terada M., Sugimura T.;
RT "Genomic sequence of hst, a transforming gene encoding a protein
RT homologous to fibroblast growth factors and the int-2-encoded
RT protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:7305-7309(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87204251; PubMed=2953031;
RA Taira M., Yoshida T., Miyagawa K., Sakamoto H., Terada M.,
RA Sugimura T.;
RT "cDNA sequence of human transforming gene hst and identification of
RT the coding sequence required for transforming activity.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2980-2984(1987).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=87301716; PubMed=2957062;
RA Dell'I Bovi P., Curatola A.M., Kern F.G., Greco A., Ittmann M.,
RA Basilico C.;
RT "An oncogene isolated by transfection of Kaposi's sarcoma DNA encodes
RT a growth factor that is a member of the FGF family.";
RL Cell 50:729-737(1987).
CC -1- FUNCTION: CAN TRANSFORM NIH 3T3 CELLS FROM A HUMAN STOMACH TUMOR
CC (HST) AND FROM KAPOSI'S SARCOMA (KS3). IT HAS A MITOGENIC
CC ACTIVITY.
CC -1- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J02986; AAB59555.1; -.
CC EMBL; M17446; AAB59473.1; -.
CC PIR; A28417; TVHUHS.
CC HSSP; P09038; 1BFF.
CC -----

DR MIM; 164980; -.
DR InterPro; IPR002209; -.
DR InterPro; IPR002348; -.
DR Pfam; PF00167; FGF; 1.
DR PRINTS; PR00262; ILIHGFG.
DR PRINTS; PR00263; HBGF-FGF.
DR PROSITE; PS00247; HBGF-FGF; 1.
KW Transforming protein; Oncogene; Growth factor; Mitogen; Signal.
FT SIGNAL 1 ? POTENTIAL.
FT CHAIN 1 ? 206 FIBROBLAST GROWTH FACTOR-4.
SQ SEQUENCE 206 AA; 22047 MW; C7FD54A0272A1569 CRC64;

Query Match
Best Local Similarity 36.9%; Score 163.5; DB 1; Length 206;
Matches 41; Conservative 19; Mismatches 46; Indels 5; Gaps 3;

QY 60 HLEIREDTGVTGGAAQSPESLLQALKPGVIOILGVKTSRFLCQPDGALYGLSHFDPE 119
DQ 95 HLQALPDGRIGGAHADTRDSLELSPVERGVVIFGVASREFVAMSSKGLYGSPTFTDE 154
QY 120 ACSFRELLEDGYNVQSEAH-GLPLHLPGN---KSPHRDPAPRGPARFLP 166
DQ 155 -CTFKEILLPNPNAYESKYKYPGMEIALSKNGTKKGNRVSPMTMKVTHFLP 204

RESULT 13
FGF5_MOUSE STANDARD; PRT; 264 AA.
AC P15656;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE FIBROBLAST GROWTH FACTOR-5 PRECURSOR (FGF-5) (HBGF-5).
GN FGF5 OR FGF-5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90201563; PubMed=2318343;
RA Hebert J.M., Basilico C., Goldfarb M., Haub O., Martin G.R.;
RT "Isolation of cDNAs encoding four mouse FGF family members and
RT characterization of their expression patterns during embryogenesis.";
RL Dev. Biol. 138:454-463(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=CS7BL/6;
RX MEDLINE=91045929; PubMed=1700424;
RA Haub O., Drucker B., Goldfarb M.;
RT "Expression of the murine fibroblast growth factor 5 gene in the
RT adult central nervous system.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8022-8026(1990).
CC -1- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M30643; AAA96698.1; -.
CC EMBL; M37823; AAB02660.1; -.
CC EMBL; M37821; AAB02660.1; JOINED.
CC EMBL; M37822; AAB02660.1; JOINED.
CC EMBL; M37821; AAB02659.1; ALT_SEQ.
CC PIR; A36207; A36207.
CC PIR; B37360; B37360.
CC HSSP; P09038; 1BFC.
CC MGD; MGI:95519; Fgf5.
```

DR InterPro; IPR002209; -
DR InterPro; IPR002348; -
DR Pfam; PF00167; FGF; 1.
DR PRINTS; PRO0262; ILIHGFG.
DR PRINTS; PRO0263; HBGFFGF.
DR PROSITE; PS00247; HBGF_FGF; 1.
DR TRANSFORMING PROTEIN; Oncogene; Growth factor; Mitogen; Signal.
FT SIGNAL 1 17
FT CHAIN 18 264
FT DOMAIN 53 59
FT CARBOHYD 108 108
FT CATABOLIC (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 264 AA; 29103 MW; F6A9C8153EE923D1 CRC64;

Query Match 14.4%; Score 161; DB 1; Length 264;
Best Local Similarity 28.6%; Pred. No. 2.3e-05;
Matches 58; Conservative 24; Mismatches 79; Indels 42; Gaps 7;

QY 34 SPPLQFGQVQRQLYTDQAQTEA-----HLEIREDTGVGAADQSPESLLQLK 84
DB 63 ASFGSGSGSEHSPQSGRRRTGSLYCRVIGFHLQIYDPGVKNGSHEASVLSILEIF 122
QY 85 ALKPGVQITLGVKTSRFLCQRPDGLALYGLSLHFDPEACSFRELLLEDGYNVYQSEAH----- 140
DB 123 AVSQGIVGIRGVFSNKLAMSKKGLHASAKFTDD-CKFRERFQENSNTYASAIHRTK 181
QY 141 -GLPLHLPNGKSPHRDPAPRG-----PARFLP-----LPGLPPAPPEPPGILAP 183
DB 182 TGREWTVALK---RGKARRGCSPRVKPQHVSTHFLPRFKSQEPFLSFVTVTVEKKKPP 238
QY 184 OPDVGSSDPLSMVGPQSGRSPS 206
DB 239 VKPKVPLSQP-----RRSPS 253

RESULT 14
FGF3_BRARE
ID FGF3_BRARE STANDARD; PRT; 256 AA.
AC P48802;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE FIBROBLAST GROWTH FACTOR-3 PRECURSOR (FGF-3) (HBGF-3).
GN FGF3 OR FGF-3
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Rasbora; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96204005; PubMed=8622866;
RA Kiefer P., Strahle U., Mason I., Dickson C.;
RT "Secretion and mitogenic activity of zebrafish FGF3 reveal intermediate properties relative to mouse and Xenopus homologues.";
RL Oncogene 12:1503-1511(1996).
CC -!- FUNCTION: POTENT MITOGEN AND TRANSFORMING AGENT (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; 248714; CAA88596.1; -
CC HSP; P09038; 2BFH.
CC ZFIN; ZDB-GENE-980526-178; fgf3.
DR InterPro; IPR002209; -
DR Pfam; PF00167; FGF; 1.
DR

DR PRINTS; PRO0262; ILIHGFG.
DR PRINTS; PRO0263; HBGFFGF.
DR PROSITE; PS00247; HBGF_FGF; 1.
KW Growth factor; Mitogen; Signal; Glycoprotein.
FT SIGNAL 1 256
FT CHAIN 1 256
FT CATABOLIC (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 256 AA; 28923 MW; CEBALC72A170BDDC CRC64;

Query Match 14.4%; Score 160.5; DB 1; Length 256;
Best Local Similarity 28.1%; Pred. No. 2.4e-05;
Matches 47; Conservative 22; Mismatches 47; Indels 51; Gaps 5;

QY 41 GGQVROLYLTDDAQOQTEAHLEIREDTGVGAADQ-----SPESLLQLKALPGVQITLGVKTS 99
DB 58 GGAPRRKKLYC-----ATKVLQIHPNGKIDGSLFENNPLSILEITAVDVGVAIKGLFSG 113
QY 100 RFLCQRPDGLALYGLSLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLPNGKSPHRDPAPR 159
DB 114 RYLAAMEKGLYASEVFNRE-CEFLERIHGELGYNVYASRHA----- 154
QY 160 GPAREFLPLGPPAPPEPPGILAPQPPDVGSSDPLSMVGPQSGRSPS 206
DB 155 -----TTQPPPTGSG-----IGSKRRASS 174

RESULT 15
FGF1_MESAU
ID FGF1_MESAU STANDARD; PRT; 155 AA.
AC P34004;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE HEPARIN-BINDING GROWTH FACTOR 1 PRECURSOR (HBGF-1) (ACIDIC FIBROBLAST GROWTH FACTOR) (AFGF).
GN FGF1 OR FGF-1.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90270291; PubMed=16933366;
RA Hall J.A., Harris M.A., Malark M., Mansson P.E., Zhou H., Harris S.E.;
RT "Characterization of the hamster DDT-1 cell aFGF/HBGF-I gene and cDNA and its modulation by steroids.";
RL J. Cell. Biochem. 43:17-26(1990).
CC -!- FUNCTION: THE HEPARIN-BINDING GROWTH FACTORS ARE ANGIOGENIC AGENTS IN VIVO AND ARE POTENT MITOGENS FOR A VARIETY OF CELL TYPES IN VITRO. THERE ARE DIFFERENCES IN THE TISSUE DISTRIBUTION AND CONCENTRATION OF THESE 2 GROWTH FACTORS.
CC -!- SUBUNIT: MONOMER.
CC -!- MISCELLANEOUS: THIS PROTEIN BINDS HEPARIN, ALTHOUGH LESS STRONGLY THAN DOES BFGF.
CC -!- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.
DR PIR; AG0721; A60721.
DR HSP; P05230; 2AXM.
DR InterPro; IPR002209; -
DR InterPro; IPR002348; -
DR Pfam; PF00167; FGF; 1.
DR PRINTS; PRO0262; ILIHGFG.
DR PRINTS; PRO0263; HBGFFGF.
DR PROSITE; PS00247; HBGF_FGF; 1.
KW Growth factor; Mitogen; Vascularization; Heparin-binding.
FT PROPEP 1 15
FT CHAIN 16 155
FT BINDING 24 28
FT BINDING 113 116
SQ SEQUENCE 155 AA; 17403 MW; 41ESEC760E412CC5 CRC64;

Query Match 13.9%; Score 155.5; DB 1; Length 155;

Best Local Similarity 35.0%; Pred. No. 3.2e-05;
Matches 48; Conservative 12; Mismatches 62; Indels 15; Gaps 5;

QY	41	GGQVRQRYLYTDDAQTEAHLEIREDTVGGAADOSPESL-LQKALKPGVIQILGVKTS	99
Db	21	GNVKKPKLLYCSNGGH---FLRLPDGTVGTRDRSDQHILQLSAESAGEVIKGTETG	77
QY	100	RELQRPDQALYGLHFDPEACSFRELLEDGYNVYQSEAHGLPLHLPCNKSPHRDPAPR	159
Db	78	QYLAMDTDGLLYGS-QTPNEECFLERLEENHYNTYTSKKHAEKNWFVGLKK--NGSCKR	134
QY	160	GPAP-----FLPLP	168
Db	135	GPRTHYGQKAILFLPLP	151

Search completed: August 6, 2001, 21:09:57
Job time: 723 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2001, 20:49:59 ; Search time 67.44 Seconds
(without alignments)
236.069 Million cell updates/sec

Title: US-09-391-861-2
Perfect score: 1116
Sequence: 1 MDSDETFEHSGLMWSVLG.....SSDPLSMVGPSQGRSPSYAS 209

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	251.5	22.5	251	JC7513	fibroblast growth
2	179	16.0	245	1 TVMST2	transforming prote
3	176.5	15.8	208	2 S20102	fibroblast growth
4	174	15.6	187	2 S23595	embryonic fibrobla
5	173	15.5	206	2 JC4268	fibroblast growth
6	170.5	15.3	208	2 S14192	fibroblast growth
7	170.5	15.3	239	1 S04742	fibroblast growth
8	168	15.1	266	2 S68144	fibroblast growth
9	165	14.8	194	2 I50710	fibroblast growth
10	163.5	14.7	192	2 S54407	embryonic fibrobla
11	163.5	14.7	206	1 TVHUS	fibroblast growth
12	161	14.4	264	2 A36207	fibroblast growth
13	160.5	14.4	256	2 JC4627	fibroblast growth
14	158	14.2	211	2 JC7353	fibroblast growth
15	157	14.1	212	3 JC7511	fibroblast growth
16	155.5	13.9	155	1 A50721	acidic fibroblast
17	154.5	13.8	155	2 S04147	acidic fibroblast
18	154.5	13.8	155	2 D37360	acidic fibroblast
19	154	13.8	208	2 S66486	fibroblast growth
20	154	13.8	208	2 A48137	fibroblast growth
21	153.5	13.8	202	1 TVMSHS	fibroblast growth
22	150	13.4	208	2 JC7082	fibroblast somatot
23	148.5	13.3	267	1 TVHUF5	acidic fibroblast
24	146.5	13.1	155	1 A33665	transforming prote
25	146	13.1	237	1 S39582	acidic fibroblast
26	144.5	12.9	155	1 GKBOA	acidic fibroblast
27	143.5	12.9	152	2 JH0476	acidic fibroblast
28	143.5	12.9	207	2 JC5941	fibroblast growth
29	142	12.7	194	2 S26049	fibroblast growth

30 142 12.7 220 2 I50588 fibroblast growth
31 141 12.6 194 2 S49501 keratinocyte growt
32 140 12.5 194 2 I48610 keratinocyte growt
33 139.5 12.5 207 2 JC5940 fibroblast growth
34 139 12.5 194 1 A36301 fibroblast growth
35 138 12.4 189 2 A48834 basic fibroblast g
36 135 12.1 155 1 A40117 basic fibroblast g
37 131.5 11.8 155 2 JH0055 acidic fibroblast
38 127.5 11.4 155 2 A60130 acidic fibroblast
39 127 11.4 164 2 S31622 basic fibroblast g
40 122 10.9 154 2 A31674 basic fibroblast g
41 122 10.9 154 2 C37360 basic fibroblast g
42 122 10.9 168 2 JG0184 fibroblast growth
43 121 10.8 137 2 I46711 fibroblast growth
44 121 10.8 146 1 S00185 basic fibroblast g
45 121 10.8 157 1 GKBOB basic fibroblast g

ALIGNMENTS

RESULT 1
JC7513
fibroblast growth factor-23 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-2001 #sequence_revision 31-Mar-2001 #text_change 31-Mar-2001
C:Accession: JC7513
R:Yamashita, T.; Yoshioka, M.; Itoh, N.
Biochem. Biophys. Res. Commun. 277, 494-498, 2000
A:Title: Identification of a novel fibroblast growth factor, FGF-23, preferentially e
A:Reference number: JC7513; MUID:20490027
A:Contents: Embryo
A:Accession: JC7513
A:Molecule type: mRNA
A:Residues: 1-251 <YAM>
A:Cross-references: DDBJ:AB037889
C:Comment: This factor, a secreted protein, acts on proximal cells and plays roles in
C:Genetics:
A:Gene: fgf-23
C:Keywords: nucleus

Query Match 22.5%; Score 251.5; DB 3; Length 251;
Best Local Similarity 32.1%; Pred. No. 1.9e-12;
Matches 76; Conservative 23; Mismatches 85; Indels 53; Gaps 9;
Qy 12 GLWVSVLGALLGACQ-----AHPIDSSPLL--QFGGQVRQRYLYTDDAQQTAEHLEIRE 66
Db 3 GTCLRLVGLVCTVCSLGTARAYPDTSPLLGNNWSLT---HLYTATA-RTSYHLQIHRD 58
Qy 67 GTVGAADOSPESLLQLKALPGVIOILGVKTSRFLCQRPDPCALYGLSLHFDPEACSFREL 126
Db 59 GHVDGTPHQTITYSALMITSEDAGSVITGAMTRFLCMLDHLGNIFGSLHFSPECKPQW 118
Qy 127 LLEDGYNVQSEAHGLPHL-----PGNKSP----- 152
Db 119 TLENGYDVLQSKHHYLSGLRAKRIQPGTNPFPFQFLARNEVLLHFIYVPRRHT 178
Qy 153 ---HNDPAPRGAPREL-PLPGLFPAP-----PEPPGILAPQPDVGVSSDPLSMVGPSQG 202
Db 179 RSAEDPPERDPLNVLKPRPRATPVVSCRELPSAEEGGP---AASDPLGVLRGRG 232

RESULT 2
TVMST2
transforming protein (int-2) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 18-Jun-1999
C:Accession: A23930; S08157
R:Moore, R.; Casey, G.; Brookes, S.; Dixon, M.; Peters, G.; Dickson, C.
EMBO J. 5, 919-924, 1986
A:Title: Sequence, topography and protein coding potential of mouse int-2: a putative
A:Reference number: A23930; MUID:86247582

Db 123 AVSOGIVGIRGVSNKFLAMSKGKLHASAKFTDD-CRFRERFOENSNTYASAIHRTEK 181
QY 141 -GLPLHLPCNKSPHRDPAPRG-----PARFLP-----LPGLP-----PAPPEPP 178
Db 182 TGREWVALNK---RGKAKRGSPRVKQHQVSTHFLPRFKOSEQPELSFTVTVPKXKPP 238
QY 179 GILAPQPPDVGSDDLSPVMGSPQGRSPS 206
Db 239 ---RPWPKPV---PLS---PSR-RSPS 255

RESULT 9
I50710
fibroblast growth factor 4 - chicken
C:Species: Gallus gallus (chicken)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C:Accession: S78506; S50858; I50710
R:Niswander, L.
submitted to the EMBL Data Library, September 1994
A:Reference number: S78506
A:Accession: S78506
A:Molecule type: DNA
A:Residues: 1-194 <NS>
A:Cross-references: EMBL:U14654; NID:G609347; PIDN:AAA58706.1; PID:G609348
R:Niswander, L.; Jeffrey, S.; Martin, G.R.; Tickle, C.
Nature 371, 609-612, 1994
A:Title: A positive feedback loop coordinates growth and patterning in the vertebrate limb
A:Reference number: I50710; MUID:95021713
A:Accession: S50858
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-136, 'Q', 'I', 'I', 'I' <NIW>
A:Cross-references: EMBL:U14654; NID:G609347; PIDN:AAA58706.1; PID:G609348
C:Genetics:
A:Gene: FGF4
C:Superfamily: fibroblast growth factor
C:Keywords: growth factor; transforming protein

Query Match 14.8%; Score 165; DB 2; Length 194;
Best Local Similarity 27.3%; Pred. No. 6.6e-06;
Matches 53; Conservative 29; Mismatches 68; Indels 44; Gaps 7;

QY 11 SGLWVSVLGALL-CACQAHPTP-----DSSP 36
Db 5 ALLPALLGLLWPCAVGRPPPGPLPGPRQRWDALFARVAPLPAERDARDGDY 64
QY 37 LQFGQGVQRVLYTDDAQOQTEAHLEIREDTGTVGGAADQSPESLLQKALKPGVQILGV 96
Db 65 LL--GYKRLRLRYCNVG--IGFHIQVLPDGRIDGHSNRYSLLETSVPERGVVSIFGV 119
QY 97 KTSRFLCQPDGALYGLHFDPEACSFRELLEDGYNVYQSEAH-GLPLHLPGN---KSP 152
Db 120 RSLGVAMNSKGLYGSTHVNDE-CFKREILLPNNNAYESKYPGMYIALSKNGRTKKG 178
QY 153 HRDPAPRGPARFLP 166
Db 179 NKVSPMTWTHFLP 192

RESULT 10
S54407
embryonic fibroblast growth factor II - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 06-Sep-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-2000
C:Accession: S54407
R:Isaacs, H.V.; Tannahill, D.; Slack, J.M.W.
Development 114, 711-720, 1992
A:Title: Expression of a novel FGF in the Xenopus embryo. A new candidate inducing factor
A:Reference number: S23595; MUID:92315916
A:Accession: S54407
A:Status: preliminary
A:Molecule type: mRNA

A:Residues: 1-192 <ISA>
A:Cross-references: EMBL:X62594; NID:G840919; PIDN:CAA44480.1; PID:G840920
C:Superfamily: fibroblast growth factor

Query Match 14.7%; Score 163.5; DB 2; Length 192;
Best Local Similarity 32.4%; Pred. No. 8.5e-06;
Matches 36; Conservative 26; Mismatches 44; Indels 5; Gaps 3;

QY 60 HLEIREDTGTVGGAADQSPESLLQKALKPGVQILGVKTSRFLCQPDGALYGLHFDPE 119
Db 81 HIQVLPDGRINGMHNENRYSLLEISPEVGVWSLYGIKSAFMVAMNAKGLYGSRYFN-E 139
QY 120 ACSFRELLEDGYNVYQSEAH-GLPLHLPGN---KSPHRDPAPRGPARFLP 166
Db 140 ECAFKEITLLPNNNAYESKYPGMYIALKNGRTKKGNNVSPMTWTHFLP 190

RESULT 11
TVHUHS
fibroblast growth factor 4 - human
N:Alternate names: heparin secretory transforming protein 1; Kaposi sarcoma oncogene;
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 18-Jun-1999
C:Accession: A28417; A29876; A29649
R:Yoshida, T.; Miyagawa, K.; Odagiri, H.; Sakamoto, H.; Little, P.F.R.; Terada, M.; S
Proc. Natl. Acad. Sci. U.S.A. 84, 7305-7309, 1987
A:Title: Genomic sequence of hst, a transforming gene encoding a protein homologous to
A:Reference number: A28417; MUID:88041096
A:Accession: A28417
A:Molecule type: DNA
A:Residues: 1-206 <YOS>
A:Cross-references: DDBJ:J02986; NID:G184430; PIDN:AAB59555.1; PID:G386788
R:Taira, M.; Yoshida, T.; Miyagawa, K.; Sakamoto, H.; Terada, M.; Sugimura, T.
Proc. Natl. Acad. Sci. U.S.A. 84, 2980-2984, 1987
A:Title: cDNA sequence of human transforming gene hst and identification of the coding
A:Reference number: A29876; MUID:87204251
A:Accession: A29876
A:Molecule type: mRNA
A:Residues: 1-206 <TAI>
A:Cross-references: GB:J02986; GB:M16338; NID:G184430; PIDN:AAB59555.1; PID:G386788
R:Deilli Bovi, P.; Curatola, A.M.; Kern, F.G.; Greco, A.; Ittmann, M.; Basilico, C.
Cell 50, 729-737, 1987
A:Title: An oncogene isolated by transfection of Kaposi's sarcoma DNA encodes a growth
A:Reference number: A29649; MUID:87301716
A:Accession: A29649
A:Molecule type: mRNA
A:Residues: 1-206 <BOV>
A:Cross-references: GB:M17446; NID:G186785; PIDN:AAA59473.1; PID:G307092
C:Comment: This protein is an oncogene for Kaposi's sarcoma. It is homologous to the
C:Genetics:
A:Gene: GDB:FGF4; HSTF1
A:Cross-references: GDB:120066; OMIM:164980
A:Map position: 11q13.3-11q13.3
A:Introns: 114/1; 148/3
C:Superfamily: fibroblast growth factor
C:Keywords: growth factor; Kaposi sarcoma; transforming protein

Query Match 14.7%; Score 163.5; DB 1; Length 206;
Best Local Similarity 36.9%; Pred. No. 9.2e-06;
Matches 41; Conservative 19; Mismatches 46; Indels 5; Gaps 3;

QY 60 HLEIREDTGTVGGAADQSPESLLQKALKPGVQILGVKTSRFLCQPDGALYGLHFDPE 119
Db 95 HLQALPDGRITGGYGAHARDLSLEISPEVGVWSIFGVSRFVAMSSKGLYGSFFTTDE 154
QY 120 ACSFRELLEDGYNVYQSEAH-GLPLHLPGN---KSPHRDPAPRGPARFLP 166
Db 155 -CTFKEILLPNNNAYESKYPGMYIALKNGRTKKGNNVSPMTWTHFLP 204

RESULT 12


```
Best Local Similarity 28.1%, Pred. No. 2e-05;
Matches 47; Conservative 22; Mismatches 47; Indels 51; Gaps 5;

Qy      41 GGQVRQLYTTDDAQQTAEHLEREDGTVGGAADQ-SPESSLQALKPGVGIQLGVKTS 99
        |||..|||.:|||.:.|||.:.|||.:.|||.:.|||.:.|||.:.|||.:.|||.:.
Db       58 GGAPRRRLYC-----AKYHLOHPNGKIDGSLLENPLSILEITADVGVVAIKGLFSG 113

Qy     100 RFLCORPDGALYLSLHFDPACSFRELLLEDGYNVYOSEAHGLPLHLPGNKSPHRDPAPR 159
        |||..|||..|||..|||..|||..|||..|||..|||..|||..|||..|||..|||..
Db      114 RYLAMNEKGRLYA SEVFENRE--CEPLRIHELGYNTYASRHHA----- 154

Qy     160 GPAREFLPGLPPAPPPEPGILAPPPDVGGSSDPLSMVGPDSGRSPS 206
        |||||.....|||||....|||.:.|||.:.|||.:.|||.:.|||.:.|||.:.|||.:.
Db     155 -----TTQPPTTGS-----IGSKRRASS 174


RESULT   14
JC7353    fibroblast growth factor-20 ~ human
C:Species: Homo sapiens (man)
C:date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 01-Dec-2000
C:Accession: JC7353
R:Kirikoshi, H.; Sagara, N.; Saitoh, T.; Tanaka, K.; Sekihara, H.; Shiokawa, K.; Katou
Biochem. Biophys. Res. Commun. 274, 337-343, 2000
```

A:Reference number: JC7353
A:Accession: JC7353
A:Molecule type: mRNA
A:Residues: 1-211 <KIR>
A:Cross-references: DDBJ:AB044277
C:Comment: This factor is involved in physiological and pathological processes, playi
C:Genetics:
A:Gene: fgf-20
A:Map position: 8p21.3-8p22
C:Superfamily: fibroblast growth factor
C:Keywords: angiogenesis; carcinogenesis

Query Match	14.2%	Score 158;	DB 2;	Length 211;
Best Local Similarity	33.1%	Pred. No. 2.5e-05;		
Matches 53;	Conservative 18;	Mismatches 65;	Indels 24;	Gaps 6;

[illegible]

RESULT 15
JC7511
fibroblast growth factor-20 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-2001 #sequence_revision 31-Mar-2001 #text_change 31-Mar-2001
C:Accession: JC7511
R:Ohmachi, S.; Watanabe, Y.; Mikami, T.; Kusu, N.; Ibi, T.; Akaike, A.; Itoh, N.
Biochem. Biophys. Res. Commun. 277, 355-360, 2000
A:Title: FGF-20, a novel neurotrophic factor, preferentially expressed in the substan
A:Reference number: JC7511; MUID:20450008
A:Contents: Brain
A:Accession: JC7511
A:Molecule type: DNA
A:Residues: 1-212 <OHM>
A:Cross-references: DDBJ:AB020021

C;Genetics:
A:Gene: fag-20

Db 40 GGAPRRKLYC-----ATKYHLQHPGSRVNGSLNSAYSILEITAVEGVVAIKLFSGR 95
QY 101 FLCQPDGALYGLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLPGNKSP--HRDPAP 158
Db 96 YLAMNKRGLYASDHNAE-CEVERIHGELGYNTYASR-----LYRTGSGPGGAQRPQGA 149
QY 159 RGP-----ARFLPLGLPPAPPEPGIILAPQPPDVGSSDPLSMVGPSPQGRSPSYAS 209
Db 150 QRPWVSVNGKGRPRRGFKTRRTQKSLFLPRVLGHKDHWMVRLQLQSSQPRAPGEGS 206

RESULT 2

US-08-867-471-11
; Sequence 11, Application US/08867471
; Patent No. 5872226
; GENERAL INFORMATION:
; APPLICANT: Nathans, Jeremy
; APPLICANT: Smallwood, Philip M.
; APPLICANT: Macke, Jennifer P.
; TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR HOMOLOGOUS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/867.471
FILING DATE: 02-JUN-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/439,725
FILING DATE: 12-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/047001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 617/678-5099
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 245 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-867-471-11

Query Match 16.0%; Score 179; DB 2; Length 245;
Best Local Similarity 28.8%; Pred. No. 3.8e-09;
Matches 51; Conservative 30; Mismatches 78; Indels 18; Gaps 5;

QY 41 GGQVRQRYLYTDDAQOEAHLEIREDTGVTGAADQSPESLLQALKKPGVIOILGVKTSR 100
Db 40 GGAPRRKLYC-----ATKYHLQHPGSRVNGSLNSAYSILEITAVEGVVAIKLFSGR 95
QY 101 FLCQPDGALYGLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLPGNKSP--HRDPAP 158
Db 96 YLAMNKRGLYASDHNAE-CEVERIHGELGYNTYASR-----LYRTGSGPGGAQRPQGA 149
QY 159 RGP-----ARFLPLGLPPAPPEPGIILAPQPPDVGSSDPLSMVGPSPQGRSPSYAS 209
Db 150 QRPWVSVNGKGRPRRGFKTRRTQKSLFLPRVLGHKDHWMVRLQLQSSQPRAPGEGS 206

RESULT 3

US-08-438-439C-7
; Sequence 7, Application US/08438439C
; Patent No. 5876967
; GENERAL INFORMATION:
; APPLICANT: Nathans, Jeremy
; APPLICANT: Smallwood, Phillip M.
; APPLICANT: Macke, Jennifer P.
; TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR HOMOLOGOUS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/438,439C
FILING DATE: May 12, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/046001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 245 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-438-439C-7

Query Match 16.0%; Score 179; DB 2; Length 245;
Best Local Similarity 28.8%; Pred. No. 3.8e-09;
Matches 51; Conservative 30; Mismatches 78; Indels 18; Gaps 5;

QY 41 GGQVRQRYLYTDDAQOEAHLEIREDTGVTGAADQSPESLLQALKKPGVIOILGVKTSR 100
Db 40 GGAPRRKLYC-----ATKYHLQHPGSRVNGSLNSAYSILEITAVEGVVAIKLFSGR 95
QY 101 FLCQPDGALYGLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLPGNKSP--HRDPAP 158
Db 96 YLAMNKRGLYASDHNAE-CEVERIHGELGYNTYASR-----LYRTGSGPGGAQRPQGA 149
QY 159 RGP-----ARFLPLGLPPAPPEPGIILAPQPPDVGSSDPLSMVGPSPQGRSPSYAS 209
Db 150 QRPWVSVNGKGRPRRGFKTRRTQKSLFLPRVLGHKDHWMVRLQLQSSQPRAPGEGS 206

RESULT 4

US-08-439-725A-14
; Sequence 14, Application US/08439725A
; Patent No. 5693775
; GENERAL INFORMATION:
; APPLICANT: Nathans, Jeremy
; APPLICANT: Smallwood, Philip M.
; APPLICANT: Macke, Jennifer P.
; TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR HOMOLOGOUS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037

;; NUMBER OF SEQUENCES: 15
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Fish & Richardson P.C.
;; STREET: 4225 Executive Square, Suite 1400
;; CITY: La Jolla
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 92037
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/439,725A
;; FILING DATE: 12-MAY-1995
;; CLASSIFICATION: 424
;;
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Haile, Lisa A.
;; REGISTRATION NUMBER: 38,347
;; REFERENCE/DOCKET NUMBER: 07265/047001
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 619/678-5070
;; TELEFAX: 617/678-5099
;;
;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 198 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;;
;; US-08-439-725A-14

Query Match 15.8%; Score 176.5; DB 1; Length 198;
Best Local Similarity 29.1%; Pred. No. 4.9e-09;
Matches 46; Conservative 27; Mismatches 52; Indels 33; Gaps 6;

QY 13 LKWSVLGALLGACQHP-----IPDS---SPLQ-----FGQV----- 44
Db 12 LKALVFLGILVGMVVPSPAGTRANNTLLDSRGWGTLLSRAGLAGETAGVNWESGYLVG 71

QY 45 --RQRYLYTDDAQTEAHLEIREDTGVTGGADQSPESLLQKALKPGVQIQLGVKTSRFL 102
Db 72 IKRRLRYCNVG--IGFHLQVLPDRISGTHEENPYSLLEISTVERGVVSLFGVRSALFV 129

QY 103 CORPDGALYGLSHFDPEACSFRELLEDGYNVYQSEAH 140
Db 130 AMNSKGRLYATPSFQ-ECKPFRETLLPNNYNAYESDLY 166

RESULT 5
US-08-441-629-12
; Sequence 12, Application US/08441629
; Patent No. 5766923
; GENERAL INFORMATION:
; APPLICANT: Kirschner, Marc W.
; APPLICANT: Kinoshita, No. 5766923iyuki
; TITLE OF INVENTION: RECEPTOR-LIGAND ASSAY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/441,629
;; FILING DATE: 15-MAY-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/279,217
;; FILING DATE: 22-JUL-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Granahan, Patricia
;; REGISTRATION NUMBER: 32,227
;; REFERENCE/DOCKET NUMBER: H095-01A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617) 861-6240
;; TELEFAX: (617) 861-9540
;; INFORMATION FOR SEQ ID NO: 12:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 198 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;;
;; US-08-441-629-12

Query Match 15.8%; Score 176.5; DB 1; Length 198;
Best Local Similarity 29.1%; Pred. No. 4.9e-09;
Matches 46; Conservative 27; Mismatches 52; Indels 33; Gaps 6;

QY 13 LKWSVLGALLGACQHP-----IPDS---SPLQ-----FGQV----- 44
Db 12 LKALVFLGILVGMVVPSPAGTRANNTLLDSRGWGTLLSRAGLAGETAGVNWESGYLVG 71

QY 45 --RQRYLYTDDAQTEAHLEIREDTGVTGGADQSPESLLQKALKPGVQIQLGVKTSRFL 102
Db 72 IKRRLRYCNVG--IGFHLQVLPDRISGTHEENPYSLLEISTVERGVVSLFGVRSALFV 129

QY 103 CORPDGALYGLSHFDPEACSFRELLEDGYNVYQSEAH 140
Db 130 AMNSKGRLYATPSFQ-ECKPFRETLLPNNYNAYESDLY 166

RESULT 6
US-08-867-471-14
; Sequence 14, Application US/08867471
; Patent No. 5872226
; GENERAL INFORMATION:
; APPLICANT: Nathans, Jeremy
; APPLICANT: Smallwood, Philip M.
; APPLICANT: Macke, Jennifer P.
; TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR HOMOLOGOUS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867,471
; FILING DATE: 02-JUN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/439,725
; FILING DATE: 12-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347

;; REFERENCE/DOCKET NUMBER: 07265/047001
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 619/678-5070
;; TELEFAX: 619/678-5099
;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 198 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-867-471-14

Query Match 15.8%; Score 176.5; DB 2; Length 198;
Best Local Similarity 29.1%; Pred. No. 4.9e-09;
Matches 46; Conservative 27; Mismatches 52; Indels 33; Gaps 6;
QY 13 LWSVLAGLLGACQAH-----IPDS---SPLIQ-----FGQV----- 44
DB 12 LWALVFLGILVGMVPSAGTRANNLLDSRGWGTLLSRAGLAGEIAGVNWESGYLVG 71
QY 45 --RORYLYTDDAQOQTEAHLEIREDTGVGGAADQSPESLLQKALKPGVYIQLGVKTSRFL 102
DB 72 IKRQRLYCNVG--IGFHLQVLPDGRISGTHEENPYSLLEISTVERGVVSLFGVRSALFV 129
QY 103 CORPDGALYGLHFDPEACSFRELLEDGYNVYQSEAH 140
DB 130 AMNSKGRLYATPSFQ-ECKFRETLLPNNYAYESDLY 166

RESULT 7
US-08-438-439C-10
; Sequence 10, Application US/08438439C
; Patent No. 5876967
; GENERAL INFORMATION:
; APPLICANT: Nathans, Jeremy
; APPLICANT: Smallwood, Phillip M.
; APPLICANT: Macke, Jennifer P.
; TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR HOMOLOGOUS
; NUMBER OF INVENTION: FACTOR-2 AND METHODS OF USE
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/438,439C
; FILING DATE: May 12, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/046001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-438-439C-10

Query Match 15.8%; Score 176.5; DB 2; Length 198;
Best Local Similarity 29.1%; Pred. No. 4.9e-09;
Matches 46; Conservative 27; Mismatches 52; Indels 33; Gaps 6;
QY 13 LWSVLAGLLGACQAH-----IPDS---SPLIQ-----FGQV----- 44
DB 12 LWALVFLGILVGMVPSAGTRANNLLDSRGWGTLLSRAGLAGEIAGVNWESGYLVG 71
QY 45 --RORYLYTDDAQOQTEAHLEIREDTGVGGAADQSPESLLQKALKPGVYIQLGVKTSRFL 102
DB 72 IKRQRLYCNVG--IGFHLQVLPDGRISGTHEENPYSLLEISTVERGVVSLFGVRSALFV 129
QY 103 CORPDGALYGLHFDPEACSFRELLEDGYNVYQSEAH 140
DB 130 AMNSKGRLYATPSFQ-ECKFRETLLPNNYAYESDLY 166

RESULT 8
US-08-438-439C-20
; Sequence 20, Application US/08438439C
; Patent No. 5876967
; GENERAL INFORMATION:
; APPLICANT: Nathans, Jeremy
; APPLICANT: Smallwood, Phillip M.
; APPLICANT: Macke, Jennifer P.
; TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR HOMOLOGOUS
; NUMBER OF INVENTION: FACTOR-2 AND METHODS OF USE
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/438,439C
; FILING DATE: May 12, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/046001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-438-439C-20

Query Match 15.8%; Score 176.5; DB 2; Length 198;
Best Local Similarity 29.1%; Pred. No. 4.9e-09;
Matches 46; Conservative 27; Mismatches 52; Indels 33; Gaps 6;
QY 13 LWSVLAGLLGACQAH-----IPDS---SPLIQ-----FGQV----- 44
DB 12 LWALVFLGILVGMVPSAGTRANNLLDSRGWGTLLSRAGLAGEIAGVNWESGYLVG 71
QY 45 --RORYLYTDDAQOQTEAHLEIREDTGVGGAADQSPESLLQKALKPGVYIQLGVKTSRFL 102

Db 72 IKQRRLXNVG--IGFHLQVLPDRIGSTHEENPYSLEISTVERGVVSLFGVRSALFV 129
QY 103 CORPDGALYGLSHFDPEACSFRELLIEDGYNVYQSEAH 140
Db 130 AMNSKRLYATPSFQ-ECKFRFTLLPNNYNAYESDLY 166

RESULT 9
US-08-718-904-15
; Sequence 15, Application US/08/18904
; Patent No. 6037329
; GENERAL INFORMATION:
; APPLICANT: Baird, J. Andrew
; APPLICANT: Chandler, Lois Ann
; APPLICANT: Sosnowski, Barbara A.
; TITLE OF INVENTION: COMPOSITIONS CONTAINING NUCLEIC ACIDS AND LIGANDS FOR THERAPE
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/718,904
; FILING DATE: 24-SEP-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6037329tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 760100.415C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
; OTHER INFORMATION: /note= "FGF-6"
; US-08-718-904-15

Query Match 15.8%; Score 176.5; DB 3; Length 198;
Best Local Similarity 29.1%; Pred. No. 4.9e-09;
Matches 46; Conservative 27; Mismatches 52; Indels 33; Gaps 6;

QY 13 LWSVVLAGLLGACQAH-----IPDS---SPLLQ-----FGQV----- 44
Db 12 LWALVFLGILVGVVSPACTRANNTLLDSRGWGTLLSRAGLAGEIAGVNWESYLVG 71

QY 45 --RQRYLTDDAQOEAHLEIREDTGVGGAADQSPESLLQLKALKPGVIQILGVKTSREL 102
Db 72 IKQRRLXNVG--IGFHLQVLPDRIGSTHEENPYSLEISTVERGVVSLFGVRSALFV 129

QY 103 CORPDGALYGLSHFDPEACSFRELLIEDGYNVYQSEAH 140
Db 130 AMNSKRLYATPSFQ-ECKFRFTLLPNNYNAYESDLY 166

RESULT 10
US-09-023-082A-14
; Sequence 14, Application US/09/023082A
; Patent No. 6077692

Query Match 15.8%; Score 176.5; DB 3; Length 198;
Best Local Similarity 29.1%; Pred. No. 4.9e-09;
Matches 46; Conservative 27; Mismatches 52; Indels 33; Gaps 6;

GENERAL INFORMATION:
APPLICANT: RUBEN, STEVEN M.
APPLICANT: JIMENEZ, PABLO
APPLICANT: DUAN, D. ROXANNE
APPLICANT: RAMPY, MARK A.
APPLICANT: MENDRICK, DONNA
APPLICANT: ZHANG, JUN
APPLICANT: NI, JIAN
APPLICANT: MOORE, PAUL A.
APPLICANT: COLEMAN, TIMOTHY A.
APPLICANT: GRUBER, JOACHIM R.
APPLICANT: DILLON, PATRICK J.
APPLICANT: GENTZ, REINER L.
TITLE OF INVENTION: KERATINOCYTE GROWTH FACTOR-2
NUMBER OF SEQUENCES: 148
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 NEW YORK AVE, NW, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/09/023,082A
FILING DATE: 13-FEB-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01790
FILING DATE: 14-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/461,195
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/023,852
FILING DATE: 13-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/039,045
FILING DATE: 28-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/862,432
FILING DATE: 23-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/910,875
FILING DATE: 13-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/055,561
FILING DATE: 13-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.0360008/EKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 198 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-09-023-082A-14

```
QY 13 LWSVVLGALLGACQHP-----IPDS---SPLLQ-----FGQV----- 44
Db 12 LWALVFLGLVGMVPSAGTRANNLLDSRGWGTLLSRAGLAGEIAGVNWESGYLVG 71
QY 45 --RQRYLYTDDAQOQTEAHLREIDGTGGAADQSPESLLQKALKPGVQIILGVKTSRFL 102
Db 72 IKRQRLYCNVG--IGFHLQVLPDGRISGTHEENPYSLLLEISTVERGVVSLFVGRSALFV 129
QY 103 CORPDGALYSLHFDPEACSFRELLLEDGYNVYQSEAH 140
Db 130 AMNSKGRLYATPSFQ-ECKFRETLPPNNYAYESDLY 166

RESULT 11
US-08-776-207-12
; Sequence 12, Application US/08776207A
; Patent No. 6080718
; GENERAL INFORMATION:
; APPLICANT: Kirschner, Marc W.
; APPLICANT: Kinoshita, No. 6080718iyuki
; TITLE OF INVENTION: Receptor-Ligand Assay
; FILE REFERENCE: HU95-01A2
; CURRENT APPLICATION NUMBER: US/08/776.207A
; EARLIER FILING DATE: 1997-06-23
; EARLIER APPLICATION NUMBER: PCT/US95/09172
; EARLIER FILING DATE: 1995-07-19
; EARLIER APPLICATION NUMBER: 08/441.629
; EARLIER FILING DATE: 1995-05-15
; EARLIER APPLICATION NUMBER: 08/279.217
; EARLIER FILING DATE: 1994-07-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Homo sapien
US-08-776-207-12

Query Match 15.8%; Score 176.5; DB 3; Length 198;
Best Local Similarity 29.1%; Pred. No. 4.9e-09;
Matches 46; Conservative 27; Mismatches 52; Indels 33; Gaps 6;

QY 13 LWSVVLGALLGACQHP-----IPDS---SPLLQ-----FGQV----- 44
Db 12 LWALVFLGLVGMVPSAGTRANNLLDSRGWGTLLSRAGLAGEIAGVNWESGYLVG 71
QY 45 --RQRYLYTDDAQOQTEAHLREIDGTGGAADQSPESLLQKALKPGVQIILGVKTSRFL 102
Db 72 IKRQRLYCNVG--IGFHLQVLPDGRISGTHEENPYSLLLEISTVERGVVSLFVGRSALFV 129
QY 103 CORPDGALYSLHFDPEACSFRELLLEDGYNVYQSEAH 140
Db 130 AMNSKGRLYATPSFQ-ECKFRETLPPNNYAYESDLY 166

RESULT 12
PCT-US95-09172-12
; Sequence 12, Application PC/TUS9509172
; GENERAL INFORMATION:
; APPLICANT: Kirschner, Marc W.
; APPLICANT: Kinoshita, Noriyuki
; TITLE OF INVENTION: RECEPTOR-LIGAND ASSAY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09172
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/279,217
; FILING DATE: 22-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/441,629
; FILING DATE: 15-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: HU95-01A PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-09172-12

Query Match 15.8%; Score 176.5; DB 5; Length 198;
Best Local Similarity 29.1%; Pred. No. 4.9e-09;
Matches 46; Conservative 27; Mismatches 52; Indels 33; Gaps 6;

QY 13 LWSVVLGALLGACQHP-----IPDS---SPLLQ-----FGQV----- 44
Db 12 LWALVFLGLVGMVPSAGTRANNLLDSRGWGTLLSRAGLAGEIAGVNWESGYLVG 71
QY 45 --RQRYLYTDDAQOQTEAHLREIDGTGGAADQSPESLLQKALKPGVQIILGVKTSRFL 102
Db 72 IKRQRLYCNVG--IGFHLQVLPDGRISGTHEENPYSLLLEISTVERGVVSLFVGRSALFV 129
QY 103 CORPDGALYSLHFDPEACSFRELLLEDGYNVYQSEAH 140
Db 130 AMNSKGRLYATPSFQ-ECKFRETLPPNNYAYESDLY 166

RESULT 13
US-08-551-171-5
; Sequence 5, Application US/08551171
; Patent No. 5679550
; GENERAL INFORMATION:
; APPLICANT: YOSHIMURA, Koji
; APPLICANT: ISHIMARU, Kaori
; APPLICANT: IGARASHI, Koichi
; APPLICANT: TERADA, Masaaki
; TITLE OF INVENTION: hst-2 MUTEN, DNA CODING FOR THE SAME AND
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESS: CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```


6:

RESULT 15
US-08-551-171-6
; Sequence 6, Application US/08551171
; Patent No. 5679550

APPLICANT: ISHIMARU, KOICHI
APPLICANT: IGARASHI, KOICHI
APPLICANT: TERADA, Masaaki
TITLE OF INVENTION: 1st-2 MUTEIN, DNA CODING FOR THE SAME AND
TITLE OF INVENTION: PREPARATION THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
ADDRESSEE: CUSHMAN
STREET: 130 Water Street
CITY: Boston
STATE: Massachusetts
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/551,171
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/196,001
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: NEUNER, George W
REGISTRATION NUMBER: 26964
REFERENCE/DOCKET NUMBER: 44146
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)523-3400
TELEFAX: (617)523-6440
TELEX: 200291 STRE UR

; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 208 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-551-171-6

Query Match 15.8%; Score 176.5; DB 1; Length 208;
Best Local Similarity 29.1%; Pred. No. 5.2e-09;
Matches 46; Conservative 27; Mismatches 52; Indels 33; Gaps 6;
QY 13 LWSVLGGLGACQHP-----IPDS---SPLLQ-----FCGQV----- 44
Db 22 LWALVFLGLVGMVVPSPAGTRANTLLDSRGWGTLLSRAGLAGETAGVNWESGYLVG 81
QY 45 --RQRYLTDDAQOQTEAHLEIREDTGVGGAQDQSPESLIQLKALKPGVQILGVKTSRFL 102
Db 82 IKRRLRYCNVG--IGFHLQVLPDGRISGTHEENPYSLLEISTVERGVVSLFGVRSALFV 139
QY 103 CORPDGALYGSLEHDPACSFRELLEDGYNVYQSEAH 140
Db 140 AMNSKGRLYATPSFQ-ECKFRETLLPNNYAYESDLY 176

Search completed: August 6, 2001, 21:05:42
Job time: 1063 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2001, 18:15:48 ; Search time 64.04 Seconds

(without alignments)
197.852 Million cell updates/sec

Title: US-09-391-861-2

Perfect score: 1116

Sequence:

1 MDSDTGFHSLWVSLAG.....SSDPLSNVGPQGRSPSYAS 209

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0601.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	803.5	72.0	210	21	AA18636
2	780	69.9	153	22	AA18636
3	273	24.5	214	20	AA18636
4	273	24.5	216	20	AA18636
5	273	24.5	216	20	AA18636
6	273	24.5	216	20	AA18636
7	273	24.5	216	20	AA18636
8	273	24.5	216	20	AA18636
9	273	24.5	216	20	AA18636
10	273	24.5	216	22	AA18636
11	273	24.5	216	22	AA18636

12	261.5	23.4	215	20	AA18636
13	229	20.5	142	22	AA18636
14	228.5	20.5	251	22	AA18636
15	206.5	18.5	213	20	AA18636
16	179	16.0	240	21	AA18636
17	176.5	15.8	198	16	AA18636
18	176.5	15.8	198	16	AA18636
19	176.5	15.8	198	16	AA18636
20	176.5	15.8	198	16	AA18636
21	176.5	15.8	198	21	AA18636
22	176.5	15.8	198	21	AA18636
23	176.5	15.8	198	22	AA18636
24	176.5	15.8	207	21	AA18636
25	176.5	15.8	208	13	AA18636
26	176.5	15.8	208	15	AA18636
27	176.5	15.8	208	22	AA18636
28	176.5	15.8	208	22	AA18636
29	176.5	15.8	208	22	AA18636
30	172	15.4	245	19	AA18636
31	170.5	15.3	239	16	AA18636
32	170.5	15.3	239	16	AA18636
33	170.5	15.3	239	19	AA18636
34	170.5	15.3	239	21	AA18636
35	170.5	15.3	239	21	AA18636
36	170.5	15.3	239	21	AA18636
37	170.5	15.3	239	22	AA18636
38	170.5	15.3	239	22	AA18636
39	170.5	15.3	239	22	AA18636
40	170.5	15.3	239	22	AA18636
41	169.5	15.2	146	15	AA18636
42	169.5	15.2	171	15	AA18636
43	169.5	15.2	175	15	AA18636
44	168	15.1	266	21	AA18636
45	163.5	14.7	140	15	AA18636

ALIGNMENTS

RESULT 1
AA18636
ID AA18636 standard; Protein; 210 AA.
XX AC
XX AA18636;
XX AC
XX 22-JAN-2001 (first entry)
XX Amino acid sequence of murine fibroblast growth factor (FGF)-21.
XX PD10-VEGFuc; gene delivery vector; eye disease; neovascular disease;
KW neurotrophic factor; anti-angiogenic factor; eye disease; glaucoma;
KW macular degeneration; diabetic retinopathy; retinitis pigmentosa;
KW inherited retinal degeneration; surgery-induced retinopathy;
KW retinal detachment; photic retinopathy; toxic retinopathy;
KW trauma-induced retinopathy; wet age related macular degeneration;
KW ARMD; retinopathy; fibroblast growth factor-20; FGF-20.
XX Mus sp.
XX WO200054813-A2.
XX 21-SEP-2000.
XX 15-MAR-2000; 2000WO-US07062.
XX 15-MAR-1999; 99US-0124460.
XX 06-JAN-2000; 2000US-0174984.
XX (CHIR) CHIRON CORP.
XX (REGC) UNIV CALIFORNIA.
XX Manning WC, Dwarki VJ, Rendahl K, Zhou S, McGee LH, Lau D;
PI Flannery JG, Miller S, Wang F, Di Polo A;

Human PRO533 prote
Human fibroblast g
Human PRO9828 prot
Human FGF-15 prote
Murine FGF-3 prote
Fibroblast growth
FGF-6, Homo sapi
Fibroblast growth
Fibroblast growth
FGF-6, SEQ ID NO:1
Human fibroblast g
FGF6 protein. Hom
Human FGF-6 protei
HST2 protein encoo
Heparin-binding se
Human FGF-6 SEQ ID
Human fibroblast g
Human fibroblast g
Fibroblast growth
Fibroblast growth
FGF-3, Homo sapi
Fibroblast growth
FGF-3, SEQ ID NO:1
Human fibroblast g
FGF3 protein. Hom
Human int-2 SEQ ID
Human fibroblast g
Human fibroblast g
Mutant heparin-bin
Mutant heparin-bin
Heparin-binding se
Rat fibroblast gro
Truncated K-RGF.

XX WPI; 1999-347718/29.
 XX Nucleic acid encoding fibroblast growth factor - 19, useful for the
 PT diagnosis, prevention and treatment of cancers
 PT
 XX Disclosure; Fig 3; 88pp; English.
 XX This invention describes a novel human fibroblast growth factor, PRO533,
 CC also known as fibroblast growth factor-19 (FGF-19). The nucleic acids,
 CC methods and PRO533 polypeptides disclosed may be used in the diagnosis
 CC and treatment of tumours and/or conditions characterized by modulation of
 CC PRO533 expression, or in the preparation of compositions for such
 CC therapies. These compositions and methods may be used in the diagnosis
 CC and treatment of neoplastic cell growth and proliferation in mammals
 CC (especially humans). The invention is based on the identification of
 CC genes that are amplified in the genome of tumour cells. Such gene
 CC amplification is expected to be associated with the over expression of
 CC the gene product and contribute to tumourgenesis and/or autocrine
 CC signalling. Accordingly, the proteins encoded by the amplified genes are
 CC believed to be useful targets for the diagnosis and/or treatment of
 CC certain cancers and may act as predictors of the prognosis for tumour
 CC treatments.
 XX Sequence 214 AA;
 SQ

Query Match 24.5%; Score 273; DB 20; Length 214;
 Best Local Similarity 36.8%; Pred. No. 3.5e-16;
 Matches 78; Conservative 27; Mismatches 81; Indels 26; Gaps 9;

QY 11 SGLWVSLAGLLGACQAHPIPDSSPLLQF--GGVQRVRLYTDDAQ-QTEAHLIEIRDG 67
 Db 12 aglwslav-agrpia-----fsdagphvhwgdpirlrlhlytsghlsscfllradg 64
 QY 68 TVGGAADQSPESLLQALKPKGVITLGVKTSRFLCQRPDGLYGLSLHFDPEACSFRELL 127
 Db 65 vvdcarqgahslleikavaltvaikgvhsrvlycmgadgkmqgllygyseedcafeeei 124
 QY 128 LEDGVNYQSEAHGLPLHLPGNKSHPRDAPRGPAREFLP-----LPGLPPAPPEPPI--- 180
 Db 125 rpdgynvyvyrsekhrlpvsissakq-rqlvknrg--flplshfplmpmvppeedlrg 180
 QY 181 ----LAPQPPDVGSDDLPMV-GPSQGRSPSY 207
 Db 181 lesdmfsspletmdmpfglvgtgleavrrpsf 212

RESULT 4
 AAY08583
 ID AAY08583 standard; Protein; 216 AA.
 XX AAY08583;
 AC
 XX 05-AUG-1999 (first entry)
 DT
 XX Human PRO533 protein fragment.
 DE
 XX PRO533; FGF-19; fibroblast growth factor; human; diagnosis; treatment;
 KW tumour; neoplastic cell growth; cell proliferation; tumourgenesis; cancer;
 KW autocrine signalling.
 XX
 OS Homo sapiens.
 XX
 PN W09927100-A1.
 XX
 XX 03-JUN-1999.
 PD
 XX 25-NOV-1998; 98WO-US25190.
 PF
 XX 21-SEP-1998; 98US-0158432.
 PR
 XX 25-NOV-1997; 97US-0066840.
 XX

PA (GETH) GENENTECH INC.
 XX Botstein D, Goddard A, Gurney AL, Hillan KJ, Lawrence DA;
 PI Roy MA;
 XX WPI; 1999-347718/29.
 DR
 XX Nucleic acid encoding fibroblast growth factor - 19, useful for the
 PT diagnosis, prevention and treatment of cancers
 PT
 XX Disclosure; Fig 11; 88pp; English.
 PS
 XX This invention describes a novel human fibroblast growth factor, PRO533,
 CC also known as fibroblast growth factor-19 (FGF-19). The nucleic acids,
 CC methods and PRO533 polypeptides disclosed may be used in the diagnosis
 CC and treatment of tumours and/or conditions characterized by modulation of
 CC PRO533 expression, or in the preparation of compositions for such
 CC therapies. These compositions and methods may be used in the diagnosis
 CC and treatment of neoplastic cell growth and proliferation in mammals
 CC (especially humans). The invention is based on the identification of
 CC genes that are amplified in the genome of tumour cells. Such gene
 CC amplification is expected to be associated with the over expression of
 CC the gene product and contribute to tumourgenesis and/or autocrine
 CC signalling. Accordingly, the proteins encoded by the amplified genes are
 CC believed to be useful targets for the diagnosis and/or treatment of
 CC certain cancers and may act as predictors of the prognosis for tumour
 CC treatments.
 XX Sequence 216 AA;
 SQ

Query Match 24.5%; Score 273; DB 20; Length 216;
 Best Local Similarity 36.8%; Pred. No. 3.6e-16;
 Matches 78; Conservative 27; Mismatches 81; Indels 26; Gaps 9;

QY 11 SGLWVSLAGLLGACQAHPIPDSSPLLQF--GGVQRVRLYTDDAQ-QTEAHLIEIRDG 67
 Db 14 aglwslav-agrpia-----fsdagphvhwgdpirlrlhlytsghlsscfllradg 66
 QY 68 TVGGAADQSPESLLQALKPKGVITLGVKTSRFLCQRPDGLYGLSLHFDPEACSFRELL 127
 Db 67 vvdcarqgahslleikavaltvaikgvhsrvlycmgadgkmqgllygyseedcafeeei 126
 QY 128 LEDGVNYQSEAHGLPLHLPGNKSHPRDAPRGPAREFLP-----LPGLPPAPPEPPI--- 180
 Db 127 rpdgynvyvyrsekhrlpvsissakq-rqlvknrg--flplshfplmpmvppeedlrg 182
 QY 181 ----LAPQPPDVGSDDLPMV-GPSQGRSPSY 207
 Db 183 lesdmfsspletmdmpfglvgtgleavrrpsf 214

RESULT 5
 AAY13353
 ID AAY13353 standard; Protein; 216 AA.
 XX AAY13353;
 AC
 XX 25-JUN-1999 (first entry)
 DT
 XX Amino acid sequence of protein PRO533.
 DE
 XX Secreted protein; transmembrane protein; human; enterocolitis;
 KW Zollinger-Ellison syndrome; gastrointestinal ulceration;
 KW congenital microvillus atrophy; skin disease; cell growth;
 KW abnormal keratinocyte differentiation; psoriasis; epithelial cancer;
 KW Parkinson's disease; Alzheimer's disease; ALS; neuropathy;
 KW fibromodulin; dermal scarring; Usher Syndrome; Atrophia areata;
 KW anti-thrombotic; wound healing; tissue repair.
 XX
 OS Homo sapiens.
 XX
 PN W09914328-A2.

XX PD 25-MAR-1999.
 XX PF 16-SEP-1998; 98WO-US19330.
 XX PR 25-NOV-1997; 97US-0066840.
 PR 17-SEP-1997; 97US-0059113.
 PR 17-SEP-1997; 97US-0059115.
 PR 17-SEP-1997; 97US-0059117.
 PR 17-SEP-1997; 97US-0059119.
 PR 17-SEP-1997; 97US-0059121.
 PR 17-SEP-1997; 97US-0059122.
 PR 17-SEP-1997; 97US-0059124.
 PR 18-SEP-1997; 97US-0059263.
 PR 18-SEP-1997; 97US-0059266.
 PR 15-OCT-1997; 97US-0062125.
 PR 17-OCT-1997; 97US-0062285.
 PR 17-OCT-1997; 97US-0062287.
 PR 21-OCT-1997; 97US-0063486.
 PR 24-OCT-1997; 97US-0062814.
 PR 24-OCT-1997; 97US-0063045.
 PR 24-OCT-1997; 97US-0063120.
 PR 24-OCT-1997; 97US-0063121.
 PR 24-OCT-1997; 97US-0063127.
 PR 24-OCT-1997; 97US-0063128.
 PR 27-OCT-1997; 97US-0063329.
 PR 27-OCT-1997; 97US-0063327.
 PR 28-OCT-1997; 97US-0063541.
 PR 28-OCT-1997; 97US-0063542.
 PR 28-OCT-1997; 97US-0063544.
 PR 28-OCT-1997; 97US-0063549.
 PR 28-OCT-1997; 97US-0063550.
 PR 29-OCT-1997; 97US-0063435.
 PR 29-OCT-1997; 97US-0063704.
 PR 29-OCT-1997; 97US-0063732.
 PR 29-OCT-1997; 97US-0063738.
 PR 29-OCT-1997; 97US-0063734.
 PR 29-OCT-1997; 97US-0064215.
 PR 29-OCT-1997; 97US-0063735.
 PR 31-OCT-1997; 97US-0063870.
 PR 31-OCT-1997; 97US-0064103.
 PR 03-NOV-1997; 97US-0064248.
 PR 07-NOV-1997; 97US-0064809.
 PR 12-NOV-1997; 97US-0065186.
 PR 17-NOV-1997; 97US-0065846.
 PR 18-NOV-1997; 97US-0065693.
 PR 21-NOV-1997; 97US-0066120.
 PR 21-NOV-1997; 97US-0066364.
 PR 24-NOV-1997; 97US-0066772.
 PR 24-NOV-1997; 97US-0066466.
 PR 24-NOV-1997; 97US-0066770.
 PR 24-NOV-1997; 97US-0066511.
 PR 24-NOV-1997; 97US-0066453.
 XX PA (GETH) GENENTECH INC.
 XX PI Chen J, Goddard A, Gurney AL, Pennica D, Wood WI, Yuan J;
 XX DR WPI; 1999-229533/19.
 XX DR N-PSDB; AAX52224.
 XX PT New isolated human genes and polypeptides used in, e.g. treatment of
 PT gastrointestinal ulceration
 XX PS Claim 12; Fig 22; 320pp; English.
 XX AA AAV13344-403 represent secreted and transmembrane human proteins.
 CC The cDNA sequences are obtained from cDNA libraries, prepared from
 CC fetal lung, fetal kidney, fetal brain, fetal liver and fetal retina.
 CC The encoded polypeptides have specific uses based on their homology to
 CC known polypeptides, e.g. PRO211 and PRO217 can be used for disorders

CC associated with the preservation and maintenance of gastrointestinal
 CC mucosa and the repair of acute and chronic mucosal lesions
 CC (e.g. enterocolitis, Zollinger-Ellison syndrome, gastrointestinal
 CC ulceration and congenital microvillus atrophy), skin diseases associated
 CC with abnormal keratinocyte differentiation (e.g. psoriasis, epithelial
 CC cancers such as lung squamous cell carcinoma of the vulva and gliomas),
 CC potent effects on cell growth and development, diseases related to
 CC growth or survival of nerve cells including Parkinson's disease,
 CC Alzheimer's disease, ALS, neuropathies or cancer. PRO265 can be used as
 CC for fibromodulin, e.g. for reducing dermal scarring. PRO264 can be used
 CC as a target for anti-tumor drugs. PRO533 may be used in the treatment
 CC of Usher Syndrome or Atrophia areata; PRO269 can be used as an
 CC anti-thrombotic agent; PRO287 polypeptides and portions may have
 CC therapeutic applications in wound healing and tissue repair; PRO317 can
 CC be used for treating problems of the kidney, uterus, endometrium, blood
 CC vessels, or related tissue, e.g. in the heart of genital tract.
 XX SQ Sequence 216 AA;
 Query Match 24.5%; Score 273; DB 20; Length 216;
 Best Local Similarity 36.8%; Pred. No. 3.6e-16;
 Matches 78; Conservative 27; Mismatches 81; Indels 26; Gaps 9;
 QY 11 SGLVSVLAGLLGACQAHPIPDSPPLQF--GGQVRQRYLYTDDAQ-QTEAHLREDDG 67
 Db 14 aglwav-agrpla-----fsdagphvhygwgdpirlrhltytsgphglsscflriradg 66
 QY 68 TVGGAADQSPESLLQLKALKPGVQILGVKTSRFLCORPDGALYGLHFDPEACSFRELL 127
 Db 67 vvdcarqgsahsilleikavalrtvaikgvhsrvylcmgadgkmgqllgyseedcafeesi 126
 QY 128 LEDGVNYVOSEAHGLPLHLPGNKSPhRDPAKGRFELP----LPGLPPAPPEPGI--- 180
 Db 127 rpdgynvyrsekhrilpvsissakq-rqlyknrg---flplshflpmpmvppeedirgh 182
 QY 181 ----LAPQPPDVGSSDPLSNV-GPSQGRSPSY 207
 Db 183 lesdmfsspletmdmpfglvtgleavirpsf 214
 AC AAY05280;
 XX 22-JUN-1999 (first entry)
 DE FGF-8 homologue PRO533.
 KW Antibody; PRO187; PRO533; PRO214; PRO240; PRO211; PRO230; PRO246;
 KW EBAF-2; inhibitor; tumour growth; cancer; EGF-like homologue;
 KW FGF-8 homologue.
 XX OS Homo sapiens.
 XX PN WO9914327-A2.
 XX PD 25-MAR-1999.
 XX PF 10-SEP-1998; 98WO-US18824.
 XX PR 25-NOV-1997; 97US-0066840.
 PR 17-SEP-1997; 97US-0059114.
 PR 18-SEP-1997; 97US-0059117.
 PR 15-OCT-1997; 97US-0059263.
 PR 17-OCT-1997; 97US-0062125.
 PR 17-OCT-1997; 97US-0062285.
 PR 17-OCT-1997; 97US-0062287.
 PR 24-OCT-1997; 97US-0062816.
 PR 29-OCT-1997; 97US-0063704.
 XX

QY 131 GYNVQSEAHGLPLHLPGNKSPHRDPAPRGPARFLPL 167

Db 132 gynyrmkhhlhifiqak-preslqddkpsnfipv 167

Search completed: August 6, 2001, 20:57:48
Job time: 9720 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 6, 2001, 21:43:02 ; Search time 1964.36 Seconds
(without alignments)
3123.105 Million cell updates/sec

Title: US-09-391-861-3
Perfect score: 649
Sequence: 1 atggaatgatgatgatctag.....ctgactcttctctgaatcta 649

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues 20456230
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	EST:*
	1: gb_est1:*
	2: gb_est2:*
	3: gb_est3:*
	4: gb_est4:*
	5: gb_est5:*
	6: gb_est6:*
	7: gb_est7:*
	8: gb_est8:*
	9: gb_est9:*
	10: gb_est10:*
	11: gb_est11:*
	12: gb_est12:*
	13: gb_est13:*
	14: gb_est14:*
	15: gb_est15:*
	16: gb_est16:*
	17: gb_est17:*
	18: gb_est18:*
	19: gb_est19:*
	20: gb_est20:*
	21: gb_est21:*
	22: gb_est22:*
	23: gb_est23:*
	24: gb_est24:*
	25: gb_est33:*
	26: gb_est34:*
	27: gb_est35:*
	28: gb_est36:*
	29: gb_est37:*
	30: gb_est38:*
	31: gb_est39:*
	32: gb_est40:*
	33: em_estba:*
	34: em_estfun:*
	35: em_esthum1:*
	36: em_esthum2:*
	37: em_esthum3:*
	38: em_esthum4:*
	39: em_esthum5:*
	40: em_esthum6:*
	41: em_esthum7:*
	42: em_esthum8:*
	43: em_esthum9:*
	44: em_esthum10:*
	45: em_esthum11:*
	46: em_esthum12:*
	47: em_esthum13:*
	48: em_esthum14:*
	49: em_esthum15:*
	50: em_esthum16:*
	51: em_esthum17:*
	52: em_esthum18:*
	53: em_esthum19:*
	54: em_esthum20:*
	55: em_esthum21:*
	56: em_esthum22:*
	57: em_esthum23:*
	58: em_esthum24:*
	59: em_esthum25:*
	60: em_esthum26:*
	61: em_esthum27:*
	62: em_esthum28:*
	63: em_estin1:*
	64: em_estin2:*
	65: em_estin3:*
	66: em_estin4:*
	67: em_estin5:*
	68: em_estomi:*
	69: em_estom2:*
	70: em_estov1:*
	71: em_estov2:*
	72: em_estpl1:*
	73: em_estpl2:*
	74: em_estpl3:*
	75: em_estpl4:*
	76: em_estpl5:*
	77: em_estpl6:*
	78: em_estpl7:*
	79: em_estpl8:*
	80: em_estpl9:*
	81: em_estpl10:*
	82: em_estro1:*
	83: em_estro2:*
	84: em_estro3:*
	85: em_estro4:*
	86: em_estro5:*
	87: em_estro6:*
	88: em_estro7:*
	89: em_estro8:*
	90: em_estro9:*
	91: em_estro10:*
	92: em_estro11:*
	93: em_estro12:*
	94: em_estro13:*
	95: em_estro14:*
	96: em_estro15:*
	97: em_estro16:*
	98: em_estro17:*
	99: em_estro18:*
	100: em_estro19:*
	101: em_estro20:*
	102: gb_est25:*
	103: gb_est26:*
	104: gb_est27:*
	105: gb_est28:*
	106: gb_est29:*
	107: gb_est30:*
	108: gb_est31:*
	109: gb_est32:*
	110: gb_est41:*
	111: gb_est42:*
	112: gb_est43:*
	113: gb_est44:*
	114: gb_est45:*
	115: gb_est46:*
	116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description
1	638.4	98.4	950	192	AK007574	Mus muscu
2	181.6	28.0	310	109	AV050323	AV050323
3	180.2	27.8	307	109	AV049138	AV049138
4	161.4	24.9	301	109	AV050161	AV050161
5	148.2	22.8	278	109	AV060326	AV060326
6	148	22.8	288	109	AV052213	AV052213
7	126.8	19.5	308	109	AV048437	AV048437
8	116.8	18.0	248	109	AV048334	AV048334
9	115.2	17.8	272	109	AV057424	AV057424
10	113.4	17.5	272	109	AV052248	AV052248
11	104.2	16.1	496	225	AQ175436	AQ175436
12	100.6	15.5	268	109	AV051279	AV051279
13	69.6	10.7	222	109	AV057345	AV057345
14	56.6	8.7	190	109	AV059644	AV059644
15	54.2	8.4	741	141	BE869144	BE869144
16	54.2	8.4	835	141	BE89616	BE89616
17	49.6	7.6	161	109	AV060716	AV060716
18	48.8	7.5	708	152	BG328684	BG328684
19	48.8	7.5	708	152	BG328684	BG328684
20	43.8	6.7	1807	192	BE785063	BE785063
21	41.8	6.4	530	175	BG308044	BG308044
22	41.2	6.3	430	30	AV397731	AV397731
23	39.6	6.1	925	219	CNS0091P	CNS0091P
24	39	6.0	454	138	BE644808	BE644808
25	37.6	5.8	213	162	BB030336	BB030336
26	37.4	5.8	477	117	AV022483	AV022483
27	36.8	5.7	522	136	BE502702	BE502702
28	36.8	5.7	550	13	AA932634	AA932634
29	36.8	5.7	641	12	AA843875	AA843875
30	36.8	5.7	840	145	BF207217	BF207217
31	36.6	5.6	311	188	R85221	R85221
32	36.6	5.6	455	119	AW652194	AW652194
33	36.6	5.6	517	12	AA846441	AA846441
34	36.6	5.6	517	136	BE476835	BE476835
35	36.6	5.6	544	32	AV662671	AV662671
36	36.6	5.6	936	219	CNS01608	CNS01608
37	36.4	5.6	342	132	BB392080	BB392080
38	36.4	5.6	519	258	L1923Y	L1923Y
39	36.4	5.6	839	175	BG325180	BG325180
40	36.2	5.6	279	125	BB092183	BB092183
41	36.2	5.6	290	127	BB187347	BB187347
42	36.2	5.6	578	32	AV691937	AV691937
43	36.2	5.6	613	32	AV694078	AV694078
44	36.2	5.6	777	32	AV693127	AV693127
45	35.8	5.5	2111	192	AK018656	AK018656

Result No.

Score

Query Match

Length

DB

ID

Description

1

638.4

98.4

950

192

AK007574

Mus muscu

2

181.6

28.0

310

109

AV050323

AV050323

3

180.2

27.8

307

109

AV049138

AV049138

4

161.4

24.9

301

109

AV050161

AV050161

5

148.2

22.8

278

109

AV060326

AV060326

6

148

22.8

288

109

AV052213

AV052213

7

126.8

19.5

308

109

AV048437

AV048437

8

116.8

18.0

248

109

AV048334

AV048334

9

115.2

17.8

272

109

AV057424

AV057424

10

113.4

17.5

272

109

AV052248

AV052248

11

104.2

16.1

496

225

AQ175436

AQ175436

12

100.6

15.5

268

109

AV051279

AV051279

13

69.6

10.7

222

109

AV057345

AV057345

14

56.6

8.7

190

109

AV

```
/note="putative"
/codon_start=1
/protein_id="BAB25115.1"
/db_xref="GI:12841203"
/translation="MEMMRGRVGTGLGVRLLAVFLGYYQVYPIPDSPSLQFGGQ
VRORYLYTDDQDEAEHLEIREDTGVVGAARHSPESLLELKALKPGVQILGVKASRF
LCQOPDCAIXSPHFDPEACSFRELLLEDGYNVYQSEAHGLPLRLPKQSPNQDATSW
GPNRFLPMPGLLHPDQAGFLPPEPPDVSSDPLSNVPELQSRSPSYAS"
BASE COUNT      204 a   278 c   241 g   227 t
ORIGIN
Query Match      98.4%; Score 638.4; DB 192; Length 950;
Best Local Similarity 99.8%; Pred. No. 9.3e-157;
Matches 639; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 atgggaatgagatcagatgtggacccctgggactgtgggtccgactctgtgctt 60
Db 185 ATGGAAATGGATGAGATCTAGAGTTGGGACCCCTGGGACTGTGGTCCGACTCTGCTGCT 244
Qy 61 gtcttcctgctgggggtctaccaagcataccccctccctgactccagccccctccctcoag 120
Db 245 GTCCTTCTGCTGGGGTCTACCAAGCATACCCATCCCTGACTCCAGCCCTCTCCAG 304
Qy 121 ttgggggtcgaagtcggcagaggtacctctacacagatgacgaccaagacactgaagcc 180
Db 305 TTTGGGGGTCAAGTCCCGCAGAGGTACCTCTACACAGATGACGACCAAGACACTGAAGCC 364
Qy 181 cactctggagatcagggagatggaacagtgttagcgcagcacacccgcagtcagaaagt 240
Db 365 CACCTGGAGATCAGGAGGATGGAACATGTGTAGCGGAGCACACCCGAGTCCAGAAAGT 424
Qy 241 ctctctggagctcaaaagccttgaagccaggggtcattcaaatcctctgggtgtcctcctct 300
Db 425 CTCTCTGGAGCTCAAAAGCCTTGAAGCCAGGGGTCAATCAAAATCCTGGGTGTCAAAGCCTCT 484
Qy 301 aggtttctttgccaacagccagatgagctctctatgtagctgcctcactttgatcctgag 360
Db 485 AGGTTTCTTTGCCAACAGCCAGATGGAGCTCTCTATGATGCGCTCACTTTGATCCTTGAG 544
Qy 361 gctcgcagcttcagagaactctgtctggaggacgtttacaatgtgtaccagctctgaagcc 420
Db 545 GCCTGCAGCTTCAGAGAAGTCTGCTGGAGGACGTTTACAATGTGTACCACTCTGAAGCC 604
Qy 421 catggcctgcccctcgctgtgctcagaaagactcccccaaacaggatgcacacatcctctgg 480
Db 605 CATGGCCTGCCCTCGCTGTGCTCAGAAAGGACTCCCCAAACAGGATGCAACACTCTCTGG 664
Qy 481 ggacctgtgccttctctccatgcagcctgtctccagagccccaagacccaagcagga 540
Db 665 GGACCTGTGCGCTTCTGCCCCATGCCAGGCTGTCTCCAGAGCCCAAGACCAAGCAGGA 724
Qy 541 ttccctgccccagagccccagatgtgggtcctctgacccccctgagcatgtgtagagcct 600
Db 725 TTCTCTGCCCCAGAGCCCCCAGATGTGGCTCTCTGACCCCTGAGCATGATGATGAGCCT 784
Qy 601 ttacagggccgaagccccagctatgctgctgactctttc 640
Db 785 TTACAGGGCCGAAGCCCCAGCTATGCTGCTGACTCTTCC 824

RESULT 2
AV050323      310 bp      mRNA      EST      22-JUN-1999
LOCUS
DEFINITION
AV050323 Mus musculus pancreas C57BL/6J adult Mus musculus cDNA
clone 1810013H18, mRNA sequence.
ACCESSION
AV050323
VERSION
AV050323.1 GI:5136095
KEYWORDS
house mouse.
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
(bases 1 to 310)
Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,
Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara
A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Kawai, J.,
Kikuchi, N., Kojima, Y., Matsuyama, T., Nitsuma, H., Oda, H., Owa, C.,
Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara
Y., Suzuki, H., Suzuki, H., Tateno, M., Tomaru, Y., Tomimaga, N.,
Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M.,
Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
RIKEN Mouse ESTs
Unpublished (1999)
Contact: Chie Owa
Genome Science Laboratory
RIKEN
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145
Fax: 81-298-36-9098
Email: genome-res@rtc.riken.go.jp
Thermostabilization and thermoactivation of thermolabile enzymes by
trehalose and its application for the synthesis of full length cDNA
(Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (http://genome.rtc.riken.go.jp) for
further details.
FEATURES
Location/Qualifiers
source
1..310
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="1810013H18"
/clone_lib="Mus musculus pancreas C57BL/6J adult"
/sex="male"
/tissue_type="pancreas"
/dev_stage="adult"
BASE COUNT      66 a   72 c   67 g   105 t
ORIGIN
Query Match      28.0%; Score 181.6; DB 109; Length 310;
Best Local Similarity 93.1%; Pred. No. 3.5e-37;
Matches 190; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Qy 437 gtctgctcagaagactccccaaaccagatgcacatcctgggagcctgtgcgcttc 496
Db 1 GTCGTCTCAGAGGACTCCCAACAGGATGCATAATCTGGGACATGTGCGATTCT 60
Qy 497 tgccatgctcaggtcgtctccagagccccaaagaccagagattcctgccccagagc 556
Db 61 TGCCCATGCCAGGCTTGTCTCCAGGAGCCCAAGACCAAGCAGGATTCCTGCCCCAGAGC 120
Qy 557 cccagatgtggctcctctgacccccctgagcatggtagagcctttacagggccgaagcc 616
Db 121 CCCAGATGTGGTCTTTTGGACCCCTGAGCATGTGTAGAGCCTTTACAGGGCCGAAGCC 180
Qy 617 ccagctatgctcctgactcttcc 640
Db 181 CCAGCTATGCTCTGAATTTTC 204

RESULT 3
AV049138      307 bp      mRNA      EST      22-JUN-1999
LOCUS
DEFINITION
AV049138 Mus musculus pancreas C57BL/6J adult Mus musculus cDNA
clone 1810008N24, mRNA sequence.
ACCESSION
AV049138
VERSION
AV049138.1 GI:5134910
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
(bases 1 to 307)
```

AUTHORS

Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K., Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara, A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M., Kawai, J., Kikuchi, N., Kojima, Y., Matsuyama, T., Niitsuma, H., Oda, H., Owa, C., Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tateno, M., Tomaru, Y., Tominaga, N., Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

TITLE

RIKEN Mouse ESTs

JOURNAL

Unpublished (1999)

COMMENT

Contact: Chie Owa

Genome Science Laboratory

RIKEN

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Tel: 81-298-36-9145

Fax: 81-298-36-9098

Email: genome-res@rtc.riken.go.jp

Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES

source

Location/Qualifiers

1..307
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="1810008N24"
/clone_lib="Mus musculus pancreas C57BL/6J adult"
/sex="male"
/tissue_type="pancreas"
/dev_stage="adult"
65 a 72 c 67 g 103 t

BASE COUNT

ORIGIN

Query Match 27.8%; Score 180.2; DB 109; Length 307;
Best Local Similarity 93.5%; Pred. No. 8.2e-37;
Matches 188; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 440 tgcctcagaaggactcccaaacaggatgcaacatctctgggacctgtgcttctctgc 499
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 TGCCTTAGAAGGAATCCCAAAACAGGATGCAACATCTGGGGACCTGTGCGCTCTG 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 500 ccattgcaggctctccacagagcccaagcaagcagattccctgcccccagagccccc 559
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 CCATGGCAGGCGCTGATTCATGAGCCCCCAAGCAAGCAGGAGTTCCTGCCCCAGAGCCCC 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 560 cagatgtggtctctctgacccctgagcagtggttagagcctttacagggccgagcccca 619
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 CAGATGTGGGATCTTTTGACCCCTGAGCATGTGAGAGCCTTTACAGGGCCGAGCCCCA 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 620 gctatgctctactcttttc 640
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 GCTATGCGTCTGTATTTTCC 201
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4

AV050161

LOCUS

DEFINITION

AV050161 Mus musculus pancreas C57BL/6J adult EST 22-JUN-1999

clone 1810012018, mRNA sequence.

ACCESSION

AV050161

VERSION

AV050161.1

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 301)

REFERENCE

AUTHORS

Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,

TITLE

RIKEN Mouse ESTs

JOURNAL

Unpublished (1999)

COMMENT

Contact: Chie Owa

Genome Science Laboratory

RIKEN

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Tel: 81-298-36-9145

Fax: 81-298-36-9098

Email: genome-res@rtc.riken.go.jp

Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES

source

Location/Qualifiers

1..301
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="1810012018"
/clone_lib="Mus musculus pancreas C57BL/6J adult"
/sex="male"
/tissue_type="pancreas"
/dev_stage="adult"
63 a 70 c 65 g 103 t

BASE COUNT

ORIGIN

Query Match 24.9%; Score 161.4; DB 109; Length 301;
Best Local Similarity 96.2%; Pred. No. 6.8e-32;
Matches 176; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 467 atgcacatctctggggacgtgtgccttctctccatcccagggcgtgtccacagagcccc 526
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 ATGCACATCTACTGGGACCTGTGCGCTTCTTGCCCATGCCAGGCCGTGTCCACGAGCCCC 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 527 aagaccaagcagattctctgccccagagccccagatgtggctctctgacccctga 586
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 AAGACCAGCAGGATTCCTGCCCCCAGAGCCCCCAGAGATGTGGGTTCCTTTGACCCCTGGA 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 587 gcatgttagagcctttacagggccgaaagccccagcagctatgcctctgactcttctgaat 646
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 GCATGTAGAGCCTTTACAGGGCCGAGGCCCGAGCTATGCGTCCCTGA-TTTTCTCTGAAT 179
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 647 cta 649
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 180 CTA 182

RESULT 5

AV060326

LOCUS

DEFINITION

AV060326 Mus musculus pancreas C57BL/6J adult EST 23-JUN-1999

clone 1810062H22, mRNA sequence.

ACCESSION

AV060326

VERSION

AV060326.1

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 278)

REFERENCE

AUTHORS

Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara

A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M., Kawai, J., Kikuchi, N., Kojima, Y., Matsuyama, T., Nitsuma, H., Oda, H., Owa, C., Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Tateno, M., Tomaru, Y., Tominaga, N., Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

TITLE JOURNAL COMMENT

RIKEN Mouse ESTs
Unpublished (1999)
Contact: Chie Owa
Genome Science Laboratory
RIKEN

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145
Fax: 81-298-36-9098

Email: genome-res@rtc.riken.go.jp

Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES source

Location/Qualifiers
1. .278
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="1810062H22"
/clone_lib="Mus musculus pancreas C57BL/6J adult"
/sex="male"
/tissue_type="pancreas"
/dev_stage="adult"
54 a 69 c 62 g 93 t

BASE COUNT ORIGIN

54 a 69 c 62 g 93 t

Query Match 22.8%; Score 148.2; DB 109; Length 278;
Best Local Similarity 95.0%; Pred. No. 1.9e-28;
Matches 153; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 480 gggactgtgccttcgccatgccaggtgctccacagagcccaagcagcagg 539
Db 12 GGGAGCTGTGCGGTTCCTCCATCCAGGCGTGTCCAGAGCCCAAGCAGG 71
QY 540 attctgtcccccagagcccccagatgtggtctctgacccctgacatgtagacc 599
Db 72 ATTCTGTCCCCCAGAGCCCCCAGATGTGGATCATTTGACCCCTGAGCATGGTAGGCC 131
QY 600 tttaggggcccgaagcccccagctatgcgtcctgactctttc 640
Db 132 TTTACAGGCGCGAGCCCGCAGCTATGCGTCTGATCTCTTC 172

RESULT 6 AV052213 LOCUS

AV052213 288 bp mRNA EST 22-JUN-1999
DEFINITION AV052213 Mus musculus pancreas C57BL/6J adult Mus musculus cDNA
clone 1810021E04, mRNA sequence.

ACCESSION AV052213 VERSION AV052213.1 KEYWORDS EST.

GI:5137985

SOURCE ORGANISM

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 288)

Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,

Akai, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara

A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M., Kawai, J.,

Kikuchi, N., Kojima, Y., Matsuyama, T., Nitsuma, H., Oda, H., Owa, C.,

Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara

Y., Suzuki, H., Tateno, M., Tomaru, Y., Tominaga, N., Watanabe, S.,

Yagame, M., Yamamura, T., Yokota, T., Yoshino, M.,

TITLE JOURNAL COMMENT

Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
RIKEN Mouse ESTs
Unpublished (1999)
Contact: Chie Owa
Genome Science Laboratory
RIKEN

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145
Fax: 81-298-36-9098

Email: genome-res@rtc.riken.go.jp

Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES source

Location/Qualifiers
1. .288
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="1810021E04"
/clone_lib="Mus musculus pancreas C57BL/6J adult"
/sex="male"
/tissue_type="pancreas"
/dev_stage="adult"
62 a 65 c 61 g 100 t

BASE COUNT ORIGIN

62 a 65 c 61 g 100 t

Query Match 22.8%; Score 148; DB 109; Length 288;
Best Local Similarity 96.8%; Pred. No. 2.2e-28;
Matches 151; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 485 ctgtgccttcctgccatgccaggtgctccacagagcccaagcagcaggattcc 544
Db 1 CTGTGCGTCTTTCCTCCATCCAGGCTGATCCAGAGCCCAAGCAGGAGGATTC 60
QY 545 tgccttcagagcccccagatgtggtctctgacccctgagcagcagcctttac 604
Db 61 TGCCCCCAGAGCCCCCAGATGTGGTTCCTCTGACCCCTGAGCATGGTAGCCCTTAC 120
QY 605 agggcgaagcccccagctatgcctcctgactctttc 640
Db 121 AGGCCGAGCCCGCAGCTATGCGTCTGATTTTTC 156

RESULT 7 AV048437 LOCUS

AV048437 308 bp mRNA EST 22-JUN-1999
DEFINITION AV048437 Mus musculus pancreas C57BL/6J adult Mus musculus cDNA
clone 1810005124, mRNA sequence.

ACCESSION AV048437 VERSION AV048437.1 KEYWORDS EST.

GI:5134209

SOURCE ORGANISM

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 308)

Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,

Akai, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara

A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M., Kawai, J.,

Kikuchi, N., Kojima, Y., Matsuyama, T., Nitsuma, H., Oda, H., Owa, C.,

Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara

Y., Suzuki, H., Tateno, M., Tomaru, Y., Tominaga, N., Watanabe, S.,

Yagame, M., Yamamura, T., Yokota, T., Yoshino, M.,

Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

RIKEN Mouse ESTs

Unpublished (1999)

Contact: Chie Owa

Genome Science Laboratory

DEFINITION					
601445130F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3849343 5', mRNA sequence.					
ACCESSION					
BE869144					
VERSION					
BE869144.1 GI:10317920					
KEYWORDS					
EST.					
SOURCE					
human.					
ORGANISM					
Homo sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE					
1 (bases 1 to 741)					
AUTHORS					
NIH-MGC http://mgc.nci.nih.gov/.					
TITLE					
National Institutes of Health, Mammalian Gene Collection (MGC)					
JOURNAL					
Unpublished (1999)					
COMMENT					
Contact: Robert Strausberg, Ph.D. Email: cgapbs-re@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM9366 row: k column: 08 High quality sequence stop: 696.					
FEATURES					
Location/Qualifiers					
1..741					
<code>/organism="Homo sapiens"</code>					
<code>/db_xref="taxon:9606"</code>					
<code>/clone="IMAGE:3849343"</code>					
<code>/clone_lib="NIH_MGC_65"</code>					
<code>/tissue_type="adenocarcinoma"</code>					
<code>/lab_host="DH10B (phage-resistant)"</code>					
<code>/note="Organ: colon; Vector: pCMV-SPORT6; Site:1; NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.8 Kb. Library constructed by Life Technologies."</code>					
BASE COUNT					
122 a 259 c 235 g 125 t					
ORIGIN					
Query Match 8.4%; Score 54.2; DB 141; Length 741;					
Best Local Similarity 48.3%; Pred. No. 0.00097;					
Matches 183; Conservative 0; Mismatches 193; Indels 3; Gaps 1					
QY	66	cctgctgggggtctaccagcatacccatccctgaactccagccccctctccagtgttg	125		
DB	166	CGTGCGGGGCCCCCTCGCCTTCGTGGAGCGGGGCCACCAGTGCACATACGGCTGGGG	225		
QY	126	gggtcaagtgcggcgagaggttaacctetacacagatgacgaaccaag---acactgaagccca	182		
DB	226	CGACCACCATCCGCTGCGCACACTGTACACCTCGGGGCCACAGGGGCTCTCCAGCTGCTT	285		
QY	183	cctggagatcaggaggatggaacagtgttaggcagcacacgcagtcacagaagtct	242		
DB	286	CCTGGCAATCCGTGCGCAGCGGCTGTGGATGTCGGCGGGGCCAGAGCGGCACAGATT	345		
QY	243	cctggagctccaagccttgaagcagggttcattcaaatcctgggtgtcacaagcctctag	302		
DB	346	GCTGGAGATCAAGGCAGTCGCTCTGCGGACCGTGCCCATCAAGGGCGTGACAGCGTGGC	405		
QY	303	gtttctttgccaacgcccagatgtagctctctatggatgcctcaacttgcactgcaggc	362		
DB	406	GTACCTCTGCATGGCGCGCGACGGCAAGATCAGGGGCTGCTCTAGTACTCGAGGAAGA	465		
QY	363	ctgcagcttcagaaactgctgctggagagcggttacaatgtgtaccagctgaagccca	422		
DB	466	CTGTGCTTTCGAGGAGGAGATCCGCCCCAGATGGGTACAAATGTGTACCGATCCGAGACA	525		
QY	423	tggcctgccctgcgtctgtg 441			
526	CGCGCTCCGGTCTCCCTG 544				

Search completed: August 6, 2001, 21:43:05
Job time: 2701 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 6, 2001, 22:28:00 ; Search time 91.85 seconds
(without alignments)
1337.647 Million cell updates/sec

Title: US-09-391-861-3

Perfect score: 649

Sequence: 1 atggaatgatgatgatctag.....ctgactcttctgaatcta 649

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*

- 1: /cgn2.6/prodata/1/ina/5A_COMB.seq.*
- 2: /cgn2.6/prodata/1/ina/5B_COMB.seq.*
- 3: /cgn2.6/prodata/1/ina/6A_COMB.seq.*
- 4: /cgn2.6/prodata/1/ina/6B_COMB.seq.*
- 5: /cgn2.6/prodata/1/ina/PCUS_COMB.seq.*
- 6: /cgn2.6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35.4	5.5	725	5	PCT-US94-04361-21
2	35.4	5.5	788	2	US-08-883-795A-35
3	35.4	5.5	823	2	US-08-785-750-1
4	35.4	5.5	1011	2	US-08-750-128-12
5	35.4	5.5	3601	3	US-09-017-631-23
6	35.4	5.5	3602	2	US-08-883-795A-33
7	35.4	5.5	3602	4	US-09-018-138-1
8	35.2	5.4	1864	1	US-08-454-720A-38
9	33.4	5.1	5975	1	US-08-404-354B-1
10	33.4	5.1	5975	1	US-08-314-083B-1
11	33.4	5.1	5975	3	US-08-435-675B-1
12	33.4	5.1	5975	3	US-08-884-599-1
13	32.8	5.1	423	1	US-08-187-780-2
14	32.8	5.1	423	1	US-08-187-780-5
15	32.8	5.1	423	2	US-08-478-485-2
16	32.8	5.1	423	2	US-08-478-485-5
17	32.8	5.1	599	6	5430019-1
18	32.8	5.1	1219	5	PCT-US93-06251-11
19	32.4	5.0	1656	3	US-09-026-958-1
20	32.2	5.0	681	5	PCT-US94-04361-22
21	32.2	5.0	681	5	PCT-US94-04361-27
22	32.2	5.0	1288	1	US-08-440-856A-9
23	32.2	5.0	5107	4	US-08-910-647-3
24	31.8	4.9	5962	6	5386025-5
25	31.6	4.9	687	5	PCT-US94-04361-25
26	31.6	4.9	3393	1	US-08-295-502-1
27	31.6	4.9	3393	5	PCT-US95-10691-1

28	31.6	4.9	4543	2	US-08-519-547A-5	Sequence 5, Appli
29	31.2	4.8	750	3	US-09-010-809-12	Sequence 12, Appl
30	31.2	4.8	1317	1	US-08-153-848-45	Sequence 45, Appl
31	31.2	4.8	1317	3	US-09-299-843A-45	Sequence 45, Appl
32	31.2	4.8	1317	5	PCT-US93-11153-45	Sequence 45, Appl
33	31.2	4.8	1469	5	PCT-US92-08258-1	Sequence 1, Appli
34	31.2	4.8	1737	1	US-08-202-056-4	Sequence 4, Appli
35	31.2	4.8	1737	1	US-08-076-093A-3	Sequence 3, Appli
36	31.2	4.8	1737	1	US-08-701-265-3	Sequence 3, Appli
37	31.2	4.8	1737	2	US-08-284-586-3	Sequence 3, Appli
38	31.2	4.8	1737	2	US-08-805-478-3	Sequence 3, Appli
39	31.2	4.8	1737	2	US-08-802-627A-3	Sequence 3, Appli
40	31.2	4.8	1737	2	US-08-801-238-3	Sequence 3, Appli
41	31.2	4.8	1737	2	US-08-801-228-3	Sequence 3, Appli
42	31.2	4.8	1737	3	US-09-104-296-3	Sequence 3, Appli
43	31.2	4.8	1737	5	PCT-US94-06380-2	Sequence 2, Appli
44	31.2	4.8	2129	1	US-08-426-819A-34	Sequence 34, Appl
45	31.2	4.8	2138	3	US-08-776-271-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
PCT-US94-04361-21
; Sequence 21, Application PC/TUS9404361
; GENERAL INFORMATION:
; APPLICANT: Brigham and Women's Hospital
; APPLICANT: 75 Francis Street
; APPLICANT: Boston, MA 02115
; APPLICANT: Bunn, H. Franklin
; APPLICANT: Wen, Danyi
; APPLICANT: Showers, Mark O.
; TITLE OF INVENTION: Erythropoietin Muteins With Enhanced
; TITLE OF INVENTION: Activity
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESS: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04361
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/049,802
; FILING DATE: 21-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 0627.336PC01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 725 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: both
; PCT-US94-04361-21

Query Match 5.5%; Score 35.4; DB 5; Length 725;
Best Local Similarity 57.8%; Pred. No. 0.33;


```

: Patent No. 5985607
:
: GENERAL INFORMATION:
:
: APPLICANT: Delcuve, Genevieve
:
: APPLICANT: Awang, Gregor
:
: TITLE OF INVENTION: Recombinant DNA Molecules and Expression
:
: TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
:
: NUMBER OF SEQUENCES: 39
:
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: BERESKIN & PARR
:
: STREET: 40 King Street West
:
: CITY: Toronto
:
: STATE: Ontario
:
: COUNTRY: Canada
:

```

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/883,795A
 FILING DATE: 27-JUN-1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Gravelle, Micheline
 REGISTRATION NUMBER: 40,261

```

> TELEPHONE: (416) 364-7311
> TELEFAX: (416) 361-1398
> INFORMATION FOR SEQ ID NO: 33:
> SEQUENCE CHARACTERISTICS:
> LENGTH: 3602 base pairs
> TYPE: nucleic acid
> STRANDEDNESS: single
> TOPOLOGY: linear
> MOLECULE TYPE: DNA (genomic)
> ORIGINAL SOURCE:
> ORGANISM: Homo sapiens
> FEATURE:
> NAME/KEY: CDS
> LOCATION: join(625..637, 1201..1346, 1605..1691, 2303..2482,
> LOCATION: 2617..2772)
> FEATURE:
> NAME/KEY: mRNA
> LOCATION: join(625..637, 1201..1346, 1605..1691, 2303..2482,
> LOCATION: 2617..2772)
>
> US-08-883-795A-33

```


TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-238-0999
TELEFAX: 619-238-0062

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 5975 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: Coding Sequence

LOCATION: 79...5700

OTHER INFORMATION:

US-08-404-354B-1

Query Match 5.1%; Score 33.4; DB 1; Length 5975;

Best Local Similarity 43.0%; Pred. No. 2.9;

Matches 163; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

QY 29 cctggagactgtgtccgaactgctgctgtcttctctgctgggggtctaccaagcat 88

Db 3605 CCCCTGGCCTTGAACGCCAAGAGCTTGAGAATCATCTCCAGCGTGAAGATGATGGTGAAG 3546

QY 89 acccctccctgactccagccctccctccagtttgggggtcaagtcccgccagaggtacc 148

Db 3545 GCACATGAGGATGCTGAGATGTGTTCACTCTCCGACTGGGTAGTGTGCTGCATG 3486

QY 149 tctacacagatgacgaccacgaactgaagccacctggagatcagggaggatggaacag 208

Db 3485 CCCAGGCAGATGCTGTGACATGATGAGGGCGAACATCAGGTATTCAAAGTAGGAGGAG 3426

QY 209 tggtaggcagacacccagctccagaagtctctgagctcaaaagccttgaagccaq 268

Db 3425 GTGACGACGTACCACACTGCTGATGTTCTTGGGATGTAGCACCGAAGTGGG 3366

QY 269 ggggtcattcaaatcctgggtgtcaagcctctaggtttcttcccaacagccagatggag 328

Db 3365 CGGCGCTTCAGGCGCATACGCACACTGGCGGTGCTTCTTGCAGCTCGCAGTCTTG 3306

QY 329 ctctctatgagtcgctcaacttctgctgagcctgcagcttcagagaactgctgctgg 388

Db 3305 TACTCCGTCCTCCCGCTGCTCCTGGAAGGTGACGATGACAAAGCCACAAAGATGTTTCATC 3246

QY 389 aggcaggttacaatgtga 407

Db 3245 ATGAAGAAGCAATGAGGA 3227

RESULT 10

US-08-314-083B-1/c

; Sequence 1, Application US/08314083B

; Patent No. 5686241

; GENERAL INFORMATION:

; APPLICANT: Ellis, Steven Bradley

; APPLICANT: Williams, Mark E.

; APPLICANT: Harpold, Michael Miller

; APPLICANT: Brenner, Robert

; APPLICANT: Schwartz, Arnold

; TITLE OF INVENTION: CALCIUM CHANNEL COMPOSITIONS AND METHODS

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Brown, Martin, Haller & McClain

; STREET: 1660 Union Street

; CITY: San Diego

; STATE: CA

; COUNTRY: USA

; ZIP: 92101-2926

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/314,083B

FILING DATE: 28-SEPT-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/914,231

FILING DATE: 13-JUL-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/603,751

FILING DATE: 08-NOV-1990

ATTORNEY/AGENT INFORMATION:

NAME: Seidman, Stephanie L

REGISTRATION NUMBER: 33,779

REFERENCE/DOCKET NUMBER: 6362-53191

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-238-0999

TELEFAX: 619-238-0062

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 5975 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: Coding Sequence

LOCATION: 79...5700

OTHER INFORMATION:

US-08-314-083B-1

Query Match

Best Local Similarity 43.0%; Pred. No. 2.9; Length 5975;

Matches 163; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

QY 29 cctggagactgtgtccgaactgctgctgtcttctctgctgggggtctaccaagcat 88

Db 3605 CCCCTGGCCTTGAACGCCAAGAGCTTGAGAATCATCTCCAGCGTGAAGATGATGGTGAAG 3546

QY 89 acccctccctgactccagccctccctccagtttgggggtccagctccagcagaggtacc 148

Db 3545 GCACATGAGGATGCTGAGATGTGTTTCATCTCTCCGACTGCTGCTGCTGCTGCTG 3486

QY 149 tctacacagatgacgaccacgaactgaagccacctggagatcagggaggatggaacag 208

Db 3485 CCCAGGCAGATGCTGTGACATGATGAGGGCGAACATCATCGGTATTCAAAGTAGGAGGAG 3426

QY 209 tggtaggcagacacccagctccagaagtctctgagctcaaaagccttgaagccaq 268

Db 3425 GTGACGACGTACCACACTGCTGATGTTCTTGGGATGTAGCACCGAAGTGGG 3366

QY 269 ggggtcattcaaatcctgggtgtcaagcctctaggtttcttcccaacagccagatggag 328

Db 3365 CGGCGCTTCAGGCGCATACTGCACACACTGGCGGTGCTTCTTGCAGCTCGCAGTCTTG 3306

QY 329 ctctctatgagtcgctcaacttctgagcctgcagcttcagagaactgctgctgg 388

Db 3305 TACTCCGTCCTCCCGCTGCTCCTGGAAGGTGACGATGACAAAGCCACAAAGATGTTTCATC 3246

QY 389 aggcaggttacaatgtga 407

Db 3245 ATGAAGAAGCAATGAGGA 3227

RESULT 11

US-08-435-675B-1/c

; Sequence 1, Application US/08435675B

; Patent No. 5710250

; GENERAL INFORMATION:

; APPLICANT: Ellis, Steven Bradley

; APPLICANT: Williams, Mark E.
 ; APPLICANT: Harpold, Michael Miller
 ; APPLICANT: Schwartz, Arnold
 ; APPLICANT: Brenner, Robert
 ; TITLE OF INVENTION: CALCIUM CHANNEL COMPOSITIONS AND METHODS
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Brown, Martin, Haller & McClain
 ; STREET: 1660 Union Street
 ; CITY: San Diego
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 92101-2926
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSEQ version 1.5
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/435,675B
 ; FILING DATE: 05-MAY-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/314,083
 ; FILING DATE: 28-SEP-1994
 ; APPLICATION NUMBER: US 07/914,231
 ; FILING DATE: 13-JUL-1992
 ; APPLICATION NUMBER: US 07/603,751
 ; FILING DATE: 08-NOV-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Seidman, Stephanie L
 ; REGISTRATION NUMBER: 33,779
 ; REFERENCE/DOCKET NUMBER: 6362-53193
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 619-238-0999
 ; TELEFAX: 619-238-0062
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 5975 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cdna
 ; FEATURE:
 ; NAME/KEY: Coding Sequence
 ; LOCATION: 79...5700
 ; OTHER INFORMATION:
 ; US-08-435-675B-1

	Query Match	5.18;	Score 33.4;	DB 1;	Length 5975;
	Best Local Similarity	43.0%;	Pred. No. 2.9;		
	Matches 163;	Conservative 0;	Mismatches 216;	Indels 0;	Gaps 0;
QY	29	ccctgggaactgtgggtccggaactgctgctggctgtctctctctctgtaggggggtctaccaagcat	88		
Db	3605	CCCTTGCCTTGAACGCCAGAGCTTGAGAAATCATCTCCAGCGTGAAGATGATGGTGAAG	3546		
QY	89	accocatccctgaactccagcccctctctccagtttgggggtccagatccgcgcagaggtacc	148		
Db	3545	GCCACATTGAGGATGTCTGAGATGTGTTTCATCTCCGACTGGTGGTAGTGTGCATG	3486		
QY	149	tctacacagatgaacgacacacactgaagcccacctggagatcaggaggatggaacag	208		
Db	3485	CCCAGGCAAGTGTGTTGAGCATGATGAGGCGCAACATCAGGTATTCAAAGTAGGAGGAG	3436		
QY	209	tggtagcgccagcacacccgcagctccagaagaatctcctcggagctcaaacgccttaagccag	268		
Db	3425	GTGACGACGTACACACCTGGTACTGGTATGGGTCTTGGGGATGTAGCACCAGGTGGG	3366		
QY	269	gggtcattccaatactctgggtgtccaaagcctctaggtttctttgccaacacagccagatggag	328		

Db 3365 CGGGCCCTTCAGGCATACACACTGACACACTGGCGCTGGTCTTTGTCACAGCTCGCAGTTCTTGT 3306
 Qy 329 ctctctatggtgatcgctcactttgatccctgagggcctgcagcttcagagaaactgctgtg 388
 Db 3305 TACTCGCTCTCCCTTGCTCCTTGGAGGTGAGATGACAAAGCCCAAGATGTTTCATC 3246
 Qy 389 aggacgggttacaatgtgta 407
 Db 3245 ATGAAGAAGGCAATGAGGA 3227
 RESULT 12
 US-08-884-599-1/c
 ; Sequence 1, Application US/08884599
 ; Patent No. 6013474
 ; GENERAL INFORMATION:
 ; APPLICANT: Ellis, Steven Bradley
 ; APPLICANT: Williams, Mark E.
 ; APPLICANT: Harpold, Michael Miller
 ; APPLICANT: Schwartz, Arnold
 ; APPLICANT: Brenner, Robert
 ; TITLE OF INVENTION: CALCIUM CHANNEL COMPOSITIONS AND METHODS
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: Brown, Martin, Haller & McClain
 ; STREET: 1660 Union Street
 ; CITY: San Diego
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 92101-2926
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSEQ Version 1.5
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/884,599
 ; FILING DATE: 27-JUNE-1997
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/314,083
 ; FILING DATE: 28-SEPT-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/914,231
 ; FILING DATE: 13-JUL-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/603,751
 ; FILING DATE: 08-NOV-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Seidman, Stephanie L.
 ; REGISTRATION NUMBER: 33,779
 ; REFERENCE/DOCKET NUMBER: 6362-53191B
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 619-238-0999
 ; TELEFAX: 619-238-0062
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 5975 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: Coding Sequence
 ; LOCATION: 79...5700
 ; OTHER INFORMATION:
 ; US-08-884-599-1

Query Match	5.1%	Score 33.4;	DB 3;	Length 5975;
Best Local Similarity	43.0%;	Pred. NO. 2.9;		
Matches 163; Conservative	0;	Mismatches 216;	Indels 0;	Gaps 0;

QY 29 ccctggagactggtggccgactgctgctgtctctctctgctgggggtctaccaagcat 88
Db 3605 CCCCTGGCCTTGACGCGCAAGAGCTTGAGAAATCATCTCCAGCGTGAAGATGATGGTGAAG 3546
QY 89 accccatccctgactcagccccctcctccagtttgggggtcaagtccggcagaggtacc 148
Db 3545 GCCACATGAGGAGTGTGAGATGTGGTTCTCTCTCCAGCTGGTGGTAGTCTGCTCATG 3486
QY 149 tctacacagatgacgacacactgaagcccccacctggagatcaggggagatggaacag 208
Db 3485 CCCAGGCAGATGGTGTGAGCATGATGAGCGCGAACAATCAGGTATTCAAAGTAGGAGGAG 3426
QY 209 tggtaggcagacacaccgcagtcacagaaagtctcctggagctcaagccttgaagccag 268
Db 3425 GTGACGACGTACACACCTGCTACTGTATGGTCTCTTGGGATGTAGCACCGAAGTGGG 3366
QY 269 ggtcattcaaatcctgggtgtcaagcctctaggtttcttcccaacagccagatggag 328
Db 3365 CGGCGCTTCAGGGCATACTGCACACACTGGCGCTGGTCTTGTCCAGCTCGCACAGTTCTTG 3306
QY 329 ctctctatggatgcctcactttgatcctgagccctgcagcttcagagaaactgctg 388
Db 3305 TACTCCGCTCTCCCTCTCTCTGAGGTGACCATGACAAAGCCACAAAGATGTTTCATC 3246
QY 389 agaacggttacaatgtgta 407
Db 3245 ATGAAGAAGGCAATGAGGA 3227

RESULT 13

US-08-187-780-2
; Sequence 2, Application US/08187780
; Patent No. 5459250
; GENERAL INFORMATION:
; APPLICANT: CLAUDIO BASILICO
; APPLICANT: DANIELA TALARICO
; TITLE OF INVENTION: MAMMALIAN GROWTH FACTOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby P.C.
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch,
; MEDIUM TYPE: 360 Kb storage
; COMPUTER: IBM or IBM-compatible
; OPERATING SYSTEM: PC/MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/187,780
; FILING DATE: January 25, 1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/901,705
; FILING DATE: June 22, 1992
; APPLICATION NUMBER: 07/806,771
; FILING DATE: December 6, 1991
; APPLICATION NUMBER: 07/177,506
; FILING DATE: April 4, 1988
; APPLICATION NUMBER: 07/062,925
; FILING DATE: June 16, 1987
; ATTORNEY/AGENT INFORMATION:
; NAME: Howard M. Frankfort
; REGISTRATION NUMBER: 32,613
; REFERENCE/DOCKET NUMBER: 5986/13586-US3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 527-7700
; TELEFAX: (212) 753-6237
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 423
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; DESCRIPTION: NO
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: This sequence can be
; OTHER INFORMATION: found on page 5, lines 3-13, in the
; OTHER INFORMATION: application, as filed.
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO: 1-423
US-08-187-780-2

Query Match 5.1%; Score 32.8; DB 1; Length 423;
Best Local Similarity 48.4%; Pred. No. 1.5;
Matches 122; Conservative 0; Mismatches 127; Indels 3; Gaps 1;
QY 162 cgaccaagacactgaagccacctggagatcaggagagatggaacagtggttaggcagc 221
Db 66 CAACGTGGGCATCGGCTTCCACCTCCAGCGCTCCCGACGCGCGCATCGGCGGCGCA 125
QY 222 acaccgcagatccagaagtctcctgtgagctcaaacgcttgaagccaggggtcattcaat 281
Db 126 CGCGGACACCCGCGACGCTGCTGGAGCTTCGCGCCCTGGAGCGGCGGTGGTGGCAT 185
QY 282 cctgggtgtcaaaagcctctaggtttcttggccaacagccagatggagctctctatggatc 341
Db 186 CTTGCGGTGGCGCAGCCGCTTCTCGTGCCATGAGCAGCAAGGCGCAAGCTCTATGGCTC 245
QY 342 gctcactttgatcctgagggctgcaggttcagagaaactgctgtgagagcggttaca 401
Db 246 GCC---CTTCTTACCAGATGATGCACGTTCAAGGAGATTCTCTTCCCAACAACATA 302
QY 402 tgtgtaccagtc 413
Db 303 CGCTACGAGTC 314
RESULT 14
US-08-187-780-5
; Sequence 5, Application US/08187780
; Patent No. 5459250
; GENERAL INFORMATION:
; APPLICANT: CLAUDIO BASILICO
; APPLICANT: DANIELA TALARICO
; TITLE OF INVENTION: MAMMALIAN GROWTH FACTOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby P.C.
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM: Diskette, 5.25 inch,
; MEDIUM TYPE: 360 Kb storage
; COMPUTER: IBM or IBM-compatible
; OPERATING SYSTEM: PC/MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/187,780
; FILING DATE: January 25, 1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/901,705
; FILING DATE: June 22, 1992
; APPLICATION NUMBER: 07/806,771
; FILING DATE: December 6, 1991
; APPLICATION NUMBER: 07/177,506
; FILING DATE: April 4, 1988
; APPLICATION NUMBER: 07/062,925
; FILING DATE: June 16, 1987
; ATTORNEY/AGENT INFORMATION:
; NAME: Howard M. Frankfort
; REGISTRATION NUMBER: 32,613
; REFERENCE/DOCKET NUMBER: 5986/13586-US3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 527-7700
; TELEFAX: (212) 753-6237
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

MEDIUM TYPE: 360 Kb storage
COMPUTER: IBM or IBM-compatible
OPERATING SYSTEM: PC/MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/187,780
FILING DATE: January 25, 1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/901,705
FILING DATE: June 22, 1992
APPLICATION NUMBER: 07/806,771
FILING DATE: December 6, 1991
APPLICATION NUMBER: 07/177,506
FILING DATE: April 4, 1988
APPLICATION NUMBER: 07/062,925
FILING DATE: June 16, 1987
ATTORNEY/AGENT INFORMATION:
NAME: Howard M. Frankfurt
REGISTRATION NUMBER: 32,613
REFERENCE/DOCKET NUMBER: 5986/13586-US3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 527-7700
TELEFAX: (212) 753-6237
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 423
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: Genomic DNA
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: This sequence
OTHER INFORMATION: corresponds to K-FGF-140 and can be
OTHER INFORMATION: found on page 11, lines 39-47 and page
OTHER INFORMATION: 12, lines 1-31, in the application, as
OTHER INFORMATION: filed.
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 1-423

US-08-187-780-5

Query Match 5.1%; Score 32.8; DB 1; Length 423;
Best Local Similarity 48.4%; Pred. No. 1.5;
Matches 122; Conservative 0; Mismatches 127; Indels 3; Gaps 1;

QY 162 cgaccacagacactgaagcccaactgagatcaggaggagatggaacagtggtaggcgcagc 221
DB 66 CAACGTGGGCATCGGCTGCCACCTCCAGCGCTCCCGACGCGCATCGCGCGCGCA 125

QY 222 acacgcagtcagaaagtctctcagctcaaaagccttgagccagcaggggtcattcaaat 281
DB 126 CGGGGACACCCCGCAGCGCTGCTGAGCTTCGCGCCCTGGAGCGGGCGGTGGAGCAT 185

QY 282 cctgggtgtcaaaagcctctagttcttttggccacagcagcagatgagctctctatggatc 341
DB 186 CTTGCGCGTGGCGACGCGGTCTTCGTGGCCATGAGCAGGCGCAAGCTCTATGGCTC 245

QY 342 gctcaactttgactcctgagggcctgcagcttcagagaactgctgctggaggacggttaca 401

DB 246 GCC---CTTCTTCCCGATGAGTGCACGTTCAAGGAGATTCTCTCCCAACAACATAAA 302
QY 402 tgtgtaccagtc 413
DB 303 CGCCTACGAGTC 314

RESULT 15
US-08-478-485-2
Sequence 2, Application US/08478485
Patent No. 5883071
GENERAL INFORMATION:
APPLICANT: CLAUDIO BASILICO
APPLICANT: DANIELA TALARICO
TITLE OF INVENTION: MAMMALIAN GROWTH FACTOR
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby P.C.
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Diskette, 3+ inch,
MEDIUM TYPE: 1.44 MB storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC/MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478,485
FILING DATE: Concurrently Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/187,780
FILING DATE: January 25, 1994
APPLICATION NUMBER: 07/901,705
FILING DATE: June 22, 1992
APPLICATION NUMBER: 07/806,771
FILING DATE: December 6, 1991
APPLICATION NUMBER: 07/177,506
FILING DATE: April 4, 1988
APPLICATION NUMBER: 07/062,925
FILING DATE: June 16, 1987
ATTORNEY/AGENT INFORMATION:
NAME: Joseph R. Robinson
REGISTRATION NUMBER: 33,448
REFERENCE/DOCKET NUMBER: 5986/13586-US6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 527-7700
TELEFAX: (212) 753-6237
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 423
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: Genomic DNA
HYPOTHETICAL: No
ANTI-SENSE: No
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: This sequence can be
OTHER INFORMATION: found on page 5, lines 3-13, in the
OTHER INFORMATION: application, as filed.
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:

VOLUME:

ISSUE:

PAGES:

DATE:

DOCUMENT NUMBER:

FILING DATE:

PUBLICATION DATE:

RELEVANT RESIDUES IN SEQ ID NO: 1-423

US-08-478-485-2

Query Match	5.18;	Score 32.8;	DB 2;	Length 423;
Best Local Similarity	48.4%;	Pred. No. 1.5;		
Matches 122;	Conservative 0;	Mismatches 127;	Indels 3;	Gaps 1;
QY 162	cgaccaagacactaagccacccctggagatcaggaggatggaacagtggtaggcgcagc	221		
Db				
QY 66	CAACGTGGGCATCGCTTCCACCTCCAGCGCTCCCGGCGCATCGGCGGCGCA	125		
Db				
QY 222	acaccgcagtcacgaagctctctggagctcaaaagccttgaagccaggggtcattcaaat	281		
Db				
QY 126	CGCGGACACCCGCGACAGCCTGCTGGAGCTCTCGCCCGTGGAGCGGGCGTGGTGAGCAT	185		
Db				
QY 282	cctgggtgtcaaaagcctctaggtttcttcttgccaacagccagatggagctctctatgatc	341		
Db				
QY 186	CTTCGGCGTGGCCAGCGGTTCTCGTGGCCATGAGCAGCAAGGGCAAGCTCTATGGCTC	245		
Db				
QY 342	gcctcacttggatcctgagcctgagcttcagagaaactgctgagagcggttataa	401		
Db				
QY 246	GCC---CTCTTCACCGATGAGTGACGTTCAAGGAGATTCTCTTCCCAACAACACTACAA	302		
Db				
QY 402	tgtgtaccagtc	413		
Db				
QY 303	CGCCTACGAGTC	314		
Db				

Search completed: August 6, 2001, 22:28:15
Job time: 4871 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 6, 2001, 22:30:48 ; Search time 163 Seconds
(without alignments)
2500.050 Million cell updates/sec

Title: US-09-391-861-3

Perfect score: 649

Sequence: 1 atggaatgatgatgatctag.....ctgactcttctcgaatcta 649

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_0601.*
1: /SIDS8/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	638.4	98.4	659	21	AAAF5631 Nucleotide sequenc
2	283.8	43.7	514	22	AAF76715 Human fibroblast g
3	68.4	10.5	996	22	AAF44266 Human PRO9828 nucl
4	54.2	8.4	2133	20	AAAX52224 Protein PRO533 CDN
5	54.2	8.4	2133	20	AAAX28430 FGF homologue PRO5
6	54.2	8.4	2137	20	AAV72455 Human PRO533 CDNA
7	54.2	8.4	2137	21	AAAC58599 Human PRO533 prote
8	54.2	8.4	2137	21	AAA30028 Human PRO533 nucle
9	54.2	8.4	2137	22	AAF58498 PRO533 coding sequ
10	54.2	8.4	2137	22	AAC87022 Nucleotide sequenc
11	54.2	8.4	2137	22	AAF72382 Human PRO533 CDNA

12	48.2	7.4	1973	21	AAAC61159 Human fibroblast g
13	48.2	7.4	2749	21	AAAC61157 Human fibroblast g
14	44.2	6.8	534	21	AAAC61160 Human fibroblast g
15	44.2	6.8	537	21	AAAC61158 Human fibroblast g
16	44	6.8	531	21	AAAC61156 Monkey fibroblast
17	44	6.8	805	21	AAAC61155 Monkey fibroblast
18	43.8	6.7	1824	20	AAV72463 Human FGF-15 DNA f
19	40.2	6.2	627	15	AAQ58992 Diospyros texana a
20	40.2	6.2	693	15	AAQ58991 Diospyros texana a
21	37.8	5.8	2426	9	AAH80751 Sense strand of Ap
22	37	5.7	3401	7	AAAG60518 Open reading frame
23	35.8	5.5	679	21	AAAF13170 Aspergillus oryzae
24	35.4	5.5	324	9	AAH81555 EPO 125 encoding e
25	35.4	5.5	582	21	AAAC46697 DNA encoding a hum
26	35.4	5.5	585	22	AAAC66882 Chimpanzee erythro
27	35.4	5.5	606	21	AAAO7253 Human EPO cDNA seq
28	35.4	5.5	616	20	AAAT77361 Polynucleotide seq
29	35.4	5.5	616	21	AAAG90393 DN encoding synthe
30	35.4	5.5	724	9	AAH81554 EPO 140B encoding
31	35.4	5.5	788	17	AAAT31529 Human erythropoiet
32	35.4	5.5	788	21	AAZ31999 Human EPO long CDN
33	35.4	5.5	823	18	AAAT64847 Human erythropoiet
34	35.4	5.5	1015	13	AAQ24282 Epo:IL-3 short, re
35	35.4	5.5	1072	13	AAQ24285 Epo:IL-3 Flex, rec
36	35.4	5.5	1158	20	AAAT25701 Human erythropoiet
37	35.4	5.5	1245	8	AAH70360 Sequence encoding
38	35.4	5.5	1255	7	AAAG60519 cDNA of clone lamb
39	35.4	5.5	1342	7	AAAG60513 Erythropoietin gen
40	35.4	5.5	1583	18	AAAT62367 Erythropoietin gen
41	35.4	5.5	1585	18	AAAT62365 Erythropoietin gen
42	35.4	5.5	1586	18	AAAT62366 Erythropoietin gen
43	35.4	5.5	1587	18	AAAT62368 Erythropoietin gen
44	35.4	5.5	1789	16	AAQ92296 Erythropoietin CDN
45	35.4	5.5	2164	21	AAA46510 DNA encoding human

ALIGNMENTS

RESULT 1

AAAT75631
ID AAA75631 standard; DNA; 659 BP.

XX AAA75631;

XX AC

XX 22-JAN-2001 (first entry)

XX DE Nucleotide sequence of murine fibroblast growth factor (FGF)-21.

XX KW pd10-VEGFuc; gene delivery vector; eye disease; neovascular disease;

XX KW neurotrophic factor; anti-angiogenic factor; eye disease; glaucoma;

XX KW macular degeneration; diabetic retinopathy; retinitis pigmentosa;

XX KW inherited retinal degeneration; surgery-induced retinopathy;

XX KW retinal detachment; photic retinopathy; toxic retinopathy;

XX KW trauma-induced retinopathy; wet age related macular degeneration;

XX KW ARMD; retinopathy; fibroblast growth factor-20; FGF-20; ss.

XX OS Mus sp.

XX FH Key Location/Qualifiers

XX CDS 14..646

XX FT /*tag= a

XX FT /product= "Fibroblast growth factor 21"

XX PN WO200054813-A2.

XX XX 21-SEP-2000.

XX XX 15-MAR-2000; 2000WO-US07062.

XX XX 15-MAR-1999; 99US-0124460.

XX PR 06-JAN-2000; 2000US-0174984.

XX XX


```

Db 221 cagaagccacctggagatcaggagatggagcgtggggcgctgctgaccagagcc 280
QY 233 cagaagctctctggagctcaagccttaagccaggggtcattcaaatcctgggtgca 292
Db 281 ccgaagctctctggagctcgaagccttggaagccggaggtattcaaatctggagtgca 340
QY 293 aagcctctaggtttctttgccacagccagatggagctctctatgagtcgctcactttg 352
Db 341 agacatccaggttctctgcccagcggccagatgggcccctgatgatgctccactttg 400
QY 353 atctcagcctgcagcttcagaaactcgtctgctggaggaacgggttacaatgtgtaccagt 412
Db 401 accctgagcctgcagctccggagctgcttcttgagcagcgatacaaatgtttaccagt 460
QY 413 ctgaagcccatggcctgcccctgcgtctgcc 443
Db 461 ccgaagcccccagcctcccgctgcacttgc 491

RESULT 3
AAF44266
ID AAF44266 standard; cDNA; 996 BP.
XX
AC AAF44266;
XX
DT 02-APR-2001 (first entry)
XX
DE Human PRO9828 nucleotide sequence SEQ ID NO:510.
XX
KW Human; secreted and transmembrane protein; PRO; cytotstatic;
KW cell death; cancer; chromosomal mapping; gene mapping; tissue typing;
KW diagnostic assay; ss.
OS Homo sapiens.
XX
PN WO200073454-A1.
XX
PD 07-DEC-2000.
XX
PF 30-MAR-2000; 2000WO-US08439.
XX
PR 02-JUN-1999; 99WO-US12252.
PR 23-JUN-1999; 99US-0141037.
PR 07-JUL-1999; 99US-0143048.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 28-JUL-1999; 99US-0146222.
PR 17-AUG-1999; 99US-0149396.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 08-OCT-1999; 99US-0158663.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 20-MAR-2000; 2000WO-US07377.
XX
PA (GETH ) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
PI Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi CJ, Gurney AL, Kijavini IJ, Napier MA, Pan J, Paoni NF;
PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;

```

```

PI Zhang Z;
XX
XX WPI: 2001-032160/04.
DR P-PSDB; AAB65297.
XX
XX PRO polynucleotides used to produce polypeptides used to target
PT bioactive molecules such as toxins, radiolabels or antibodies, to
PT specific cells, to cause targeted cell death -
XX
XX Claim 2; Fig 323; 935pp; English.
XX
XX The present invention describes human secreted and transmembrane PRO
CC proteins. The PRO proteins have cytostatic activity. The PRO proteins
CC can be used for targeted delivery of bioactive molecules, such as
CC toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide
CC sequences, and their fragments, can be used as hybridisation probes, in
CC chromosomal and gene mapping, and in the generation of anti-sense RNA
CC and DNA. They may also be used to produce transgenic animals which are
CC nucleotide and protein sequence can be used for tissue typing and in
CC treating cancer. Anti-PRO antibodies can be used in diagnostic assays.
CC AAF44270 to AAF44470 represent PCR primers and hybridisation probes used
CC in the isolation of human PRO sequences. AAF4087 to AAF44269 and
CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein
CC sequences given in the exemplification of the present invention.
XX
SQ Sequence 996 BP; 222 A; 320 C; 260 G; 194 T; 0 other;

Query Match 10.5%; Score 68.4; DB 22; Length 996;
Best Local Similarity 53.8%; Pred. No. 1.9e-09;
Matches 141; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 180 ccactggagatcaggagagatggaacagtggttaggcgcagcacacccagtcagaaag 239
Db 384 ccactcgagatccacaagaatggcctggtggtggcgccaccatcctctacag 443
QY 240 tctcctggagctcaagccttggaagccagggttcattcaaatcctggtgtcaagcctc 299
Db 444 tgcctgatgatcagatcagagagatgctgcttgggtgattacaggtgtgatgagcag 503
QY 300 taggtttctttggccacagccagatggagctctctatggtgcctcactttgatctga 359
Db 504 aagatacctctgcattgatttcagaggcaacatttttggatcacactatttcgaccgga 563
QY 360 ggcctgcagcttcagaaactgctgctggagacggttacaaatgtgtaccagttgaagc 419
Db 564 gaactgcaggttccacaccagacgctggaaacgggtacgaogtctaccactctctca 623
QY 420 ccattggcctgcccctgcgtctg 441
Db 624 gataccttctcgtgctgctg 645

RESULT 4
AA552224
ID AAX52224 standard; DNA; 2133 BP.
XX
AC AAX52224;
XX
XX 25-JUN-1999 (first entry)
XX
XX Protein PRO533 cDNA clone DNA49435-1219.
XX
XX Secreted protein; transmembrane protein; human; enterocolitis;
KW Zollinger-Ellison syndrome; gastrointestinal ulceration;
KW congenital microvillus atrophy; skin disease; cell growth;
KW abnormal keratinocyte differentiation; psoriasis; epithelial cancer;
KW Parkinson's disease; Alzheimer's disease; ALS; neuropathy;
KW fibromodulin; dermal scarring; Usher Syndrome; Atrophila areata;
KW anti-thrombotic; wound healing; tissue repair; ss.
XX
OS Homo sapiens.

```



```

PD XX 25-MAR-1999.
PF XX 10-SEP-1998; 98WO-US18824.
PR XX 25-NOV-1997; 97US-0066840.
PR XX 17-SEP-1997; 97US-0059114.
PR XX 17-SEP-1997; 97US-0059117.
PR XX 18-SEP-1997; 97US-0059263.
PR XX 15-OCT-1997; 97US-0062125.
PR XX 17-OCT-1997; 97US-0062285.
PR XX 17-OCT-1997; 97US-0062287.
PR XX 24-OCT-1997; 97US-0062816.
PR XX 23-OCT-1997; 97US-0063704.
PA XX (GETH ) GENENTECH INC.
XX XX Botstein D, Goddard A, Gurney A, Hillan K, Lawrence DA;
PI XX Roy M, Wood WI;
XX XX WPI; 1999-229532/19.
DR XX P-PSDB; AAY05280.
XX XX Antibodies against specific proteins overexpressed in tumours
XX XX Example 1; Fig 5; 130pp; English.
XX XX This sequence encodes the FGF homologue PRO533.
CC XX The invention relates to antibodies (Ab) that bind to any of the
CC XX polypeptides (I) designated PRO187; PRO533; PRO214; PRO240; PRO211;
CC XX PRO230; PRO261; PRO246 or EBAF-2. The Ab, or other agents that inhibit
CC XX expression and/or activity of (I) are used: (i) to inhibit growth of
CC XX tumours; and (ii) as diagnostic/prognostic reagents for detection or
CC XX quantification of (I) in cells or tissues, by standard immunoassays, with
CC XX overexpression being indicative of cancer. For therapeutic use, the Ab
CC XX may be conjugated to a toxin, chemotherapeutic agent or radioisotope.
CC XX Genes expressing (I), many of which are growth factor homologues, are
CC XX overexpressed in some cases of cancer.
XX XX
SQ Sequence 2133 BP; 422 A; 641 C; 598 G; 472 T; 0 other;

Query Match 8.4%; Score 54.2; DB 20; Length 2133;
Best Local Similarity 48.3%; Pred. No. 2e-05;
Matches 183; Conservative 0; Mismatches 193; Indels 3; Gaps 1;

QY 66 cctgctgggggtctaccagcaccatccctgactcagcccccctcctcagttgg 125
DB 66 cctgctgggggtctaccagcaccatccctgactcagcccccctcctcagttgg 125
QY 126 ggggtcaagtcgagaggtacacctacacagatgacgaccaag---acactgaagccca 182
DB 126 ggggtcaagtcgagaggtacacctacacagatgacgaccaag---acactgaagccca 182
QY 580 cgaccccatcgctcgagcactgtacacctcgcccccacgggtctcctcagctgtt 639
DB 580 cgaccccatcgctcgagcactgtacacctcgcccccacgggtctcctcagctgtt 639
QY 183 cctggagatcaggagaggtggaacagtggtgtaggcgcagcacaccgcagtcagaaagtct 242
DB 183 cctggagatcaggagaggtggaacagtggtgtaggcgcagcacaccgcagtcagaaagtct 242
DB 640 cctgcgcacgtgcgcagcgcgtgtagtcgactgcgcgggggcccagagcgcgacagttt 699

RESULT 6
AAV72455 standard; cDNA; 2137 BP.
XX AAV72455;
XX AAV72455;
XX 05-AUG-1999 (first entry)
XX Human PRO533 cDNA.
XX PRO533; FGF-19; fibroblast growth factor; human; diagnosis; treatment;
KW tumour; neoplastic cell growth; cell proliferation; tumorigenesis; cancer;
KW autocrine signalling; ss.
XX Homo sapiens.
XX WO9927100-A1.
XX 03-JUN-1999.
XX 25-NOV-1998; 98WO-US25190.
XX 21-SEP-1998; 98US-0158432.
XX 25-NOV-1997; 97US-0066840.
XX (GETH ) GENENTECH INC.
XX Botstein D, Goddard A, Gurney AL, Hillan KJ, Lawrence DA;
PI Roy MA;
XX WPI; 1999-347718/29.
DR P-PSDB; AAY08580.
XX Nucleic acid encoding fibroblast growth factor - 19, useful for the
PT diagnosis, prevention and treatment of cancers
XX Claim 4; Fig 2; 88pp; English.
XX This invention describes a novel human fibroblast growth factor, PRO533,
CC also known as fibroblast growth factor-19 (FGF-19). The nucleic acids,
CC methods and PRO533 polypeptides disclosed may be used in the diagnosis
CC and treatment of tumours and/or conditions characterized by modulation of
CC PRO533 expression, or in the preparation of compositions for such
CC therapies. These compositions and methods may be used in the diagnosis
CC and treatment of neoplastic cell growth and proliferation in mammals
CC (especially humans). The invention is based on the identification of
CC genes that are amplified in the genome of tumour cells. Such gene
CC amplification is expected to be associated with the over expression of
CC the gene product and contribute to tumourigenesis and/or autocrine
CC signalling. Accordingly, the proteins encoded by the amplified genes are
CC believed to be useful targets for the diagnosis and/or treatment of
CC certain cancers and may act as predictors of the prognosis for tumour
CC treatments.
XX Sequence 2137 BP; 422 A; 648 C; 598 G; 469 T; 0 other;

Query Match 8.4%; Score 54.2; DB 20; Length 2137;
Best Local Similarity 48.3%; Pred. No. 2e-05;
Matches 183; Conservative 0; Mismatches 193; Indels 3; Gaps 1;

QY 66 cctgctgggggtctaccagcaccatccctgactcagcccccctcctcagttgg 125
DB 66 cctgctgggggtctaccagcaccatccctgactcagcccccctcctcagttgg 125
QY 520 cgtggccggggcccccctcctcgtcgagcggggggcccccacgtgcactacgctggg 579
DB 520 cgtggccggggcccccctcctcgtcgagcggggggcccccacgtgcactacgctggg 579
QY 126 ggggtcaagtcgagaggtacacctacacagatgacgaccaag---acactgaagccca 182
DB 126 ggggtcaagtcgagaggtacacctacacagatgacgaccaag---acactgaagccca 182
QY 580 cgaccccatcgctcgagcactgtacacctcgcccccacgggtctcctcagctgtt 639
DB 580 cgaccccatcgctcgagcactgtacacctcgcccccacgggtctcctcagctgtt 639
QY 183 cctggagatcaggagaggtggaacagtggtgtaggcgcagcacaccgcagtcagaaagtct 242
DB 183 cctggagatcaggagaggtggaacagtggtgtaggcgcagcacaccgcagtcagaaagtct 242
DB 640 cctgcgcacgtgcgcagcgcgtgtagtcgactgcgcgggggcccagagcgcgacagttt 699

```



```
Db 700 gctggagatcaaggcagctcgtcgtgaccgtgcccataaggcgtgcacagcgtgcg 759
Qy 303 gttctttgccaacgacagatggagctctctatgatcgctcacttctgacctgaggc 362
Db 760 gtactctgcatggtggcgcgcagcgcgaagatgcagggggtgttcaagtactcggaggaaga 819
Qy 363 ctgcagcttcagagaactcgtcgtgagggacggtttacaatgtgtaccagttctgaagccca 422
Db 820 ctgtgctttcgaggagagatcgcccgatggctacaatgtgtaccgatccgagaagca 879
Qy 423 tggcctgcccctgcgtctg 441
Db 880 ccgcctccggtctccctg 898

RESULT 8
AA30028
ID AAA30028 standard; cDNA; 2137 BP.
XX
AC AAA30028;
XX
DT 09-AUG-2000 (first entry)
XX
DE Human PRO533 nucleotide sequence.
XX
KW Antibody; PRO187; PRO533; PRO214; PRO240; PRO211; PRO230; PRO261; PRO246;
KW PRO317; tumour growth inhibitor; cancer; diagnosis; treatment; human;
KW cell growth; proliferation; fibroblast growth factor; ADEPT;
KW antibody dependent enzyme mediated prodrug therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200015666-A2.
XX
PD 23-MAR-2000.
XX
PF 08-SEP-1999; 99WO-US20594.
XX
PR 10-SEP-1998; 98US-0099803.
XX
PR 10-SEP-1998; 98WO-US18824.
XX
PA (GETH ) GENENTECH INC.
XX
PI Goddard A, Gurney AL, Hillan KJ, Roy MA, Wood WI, Botstein D;
XX
DR WPI; 2000-271386/23.
XX
DR P-PSDB; AAY88568.
XX
PT New isolated antibodies which bind to specific polypeptides used for
PS diagnosis and treatment of neoplastic cell growth and proliferation -
PS Example 2; Fig 3; 200pp; English.
XX
CC This sequence represents a human PRO533 nucleotide sequence. PRO533
CC shares sequence homology with the fibroblast growth factor. The
CC invention relates to isolated antibodies which bind to a polypeptide.
CC The "PRO" polypeptides are encoded by genes which are over expressed in
CC the genome of tumour cells. Vectors and host cells comprising the nucleic
CC acid encoding the antibodies are used in the production of the
CC antibodies. The antibodies and nucleic acids encoding them are used for
CC diagnosing a tumour in a mammal. The antibodies are used for inhibiting
CC the growth of tumour cells and identifying compounds that inhibit a
CC biological or immunological activity of and/or expression of a PRO187,
CC PRO533, PRO214, PRO240, PRO211, PRO230, PRO261, PRO246 or PRO317
CC mediated prodrug therapy (ADEPT) by conjugating the antibody to a
CC prodrug-activating enzyme which converts a prodrug to an anti-cancer
CC drug. The antibodies can be fluorescently labelled and monitored by light
CC microscopy, flow cytometry or fluorimetry for diagnosis and prognosis of
CC tumours.
XX
SQ Sequence 2137 BP; 422 A; 648 C; 598 G; 469 T; 0 other;
```

```
Query Match 8.4%; Score 54.2; DB 21; Length 2137;
Best Local Similarity 48.3%; Pred. No. 2e-05;
Matches 183; Conservative 0; Mismatches 193; Indels 3; Gaps 1;

Qy 66 cctgctgggggtaccacaatacccccctcctgactccagccccctctcctcagttgg 125
Db 520 cgtggcggggcggccccctcctctctcgacgcgggggcccccagctgcactacggctggg 579
Qy 126 gggccaagtccgcagaggtacctctacacagatgacgacccaag---acaactgaagccca 182
Db 580 cgaccctccgctcggcgacctgtacacctcggccccccacggcctctcccagctgctt 639
Qy 183 cctggagatcaggagagatggaacagtggtaggcgacgacacccgcagtcacagaaagtct 242
Db 640 cctgcgcatccgtgcgacgagcgtcgtggactgcgcggggccagagcgcacagttt 699
Qy 243 cctggagctcaaaagccttgaagccaggggtcattcaaatcctcctgggtgtcaaaagcctctag 302
Db 700 gctggagatcaaggcagctcgtcgtcgacccgtggccatcaaggcgtgcacagcgtgcg 759
Qy 303 gttctttgccaacgacagatggagctctctatgatcgctcacttctgacctgaggc 362
Db 760 gtactctgcatggtggcgcgcagcgcgaagatgcagggggtcctcagtaactcggaggaaga 819
Qy 363 ctgcagcttcagagaactcgtcgtggagagcgtttacaatgtgtaccagttctgaagccca 422
Db 820 ctgtgctttcgaggagagatccgcagatcgcgccagatggctacaatgtgtaccgatccgagaagca 879
Qy 423 tggcctgcccctgcgtctg 441
Db 880 ccgcctccggtctccctg 898

RESULT 9
AAF58498
ID AAF58498 standard; cDNA; 2137 BP.
XX
AC AAF58498;
XX
DT 27-APR-2001 (first entry)
XX
DE PRO533 coding sequence.
XX
KW Cytostatic; PRO protein; tumour; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200105836-A1.
XX
PD 25-JAN-2001.
XX
PF 20-DEC-1999; 99WO-US30999.
XX
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 02-DEC-1999; 99WO-US28564.
XX
(GETH ) GENENTECH INC.
XX
PI Botstein D, Goddard A, Gurney AL, Hillan KJ, Roy MA, Wood WI;
XX
XX WPI; 2001-091968/10.
XX
XX P-PSDB; AAB68593.
XX
PT New antibody that binds to a PRO polypeptide, e.g. PRO187 and PRO533,
```

PT useful for diagnosing and treating cancers -

XX Claim 50; Fig 3; 196pp; English.

XX The present invention relates to PRO proteins and coding sequences. The
CC present sequence is the coding sequence for one such PRO protein.
CC It was found that the PRO genes are amplified in the genome of tumour
CC cells. The gene amplification is expected to be associated with the
CC overexpression of the gene product and contributes to tumorigenesis.
CC Therefore, antagonists of PRO proteins are useful for the treatment of
CC benign or malignant tumours, leukaemias, lymphoid malignancies and other
CC disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,
CC epithelial, inflammatory and immunologic disorders.

XX Sequence 2137 BP; 422 A; 648 C; 598 G; 469 T; 0 other;

Query Match 8.4%; Score 54.2; DB 22; Length 2137;
Best Local Similarity 48.3%; Pred. NO. 2e-05;
Matches 183; Conservative 0; Mismatches 193; Indels 3; Gaps 1;

QY 66 cctgctgggggtctaccaagcattaccatccctgaactccagccccctccagtttgg 125
DB 520 cgtggccggggcccccctcgtcttcgagcggggcccccacgtgcactacggctggg 579
QY 126 ggttaagtcggcagaggtacctctacacagatgacgaccaag---acactgaagccca 182
DB 580 cgaccccatccgctcgccacgtgtacacctccgccccacgggtctccagctgtt 639
QY 183 cctgagatcagggaggtggaacagtggtgagcgcagcacccagtcacagaaagtct 242
DB 640 cctgcgcatccgtgcgacgctgtgactgcgctggggggcgagcgcgacagttt 699
QY 243 cctgagctcaagccttgaagccgggtgattcaaatcctgggtgtcaagcctctag 302
DB 700 gctggagatcaaggcagctcgtcggaccggtggccatcaaggcgcgacagcgtgcg 759
QY 303 gttctttgccaacgacagatggagctctctatgctgcctcacttctgactcagtcg 362
DB 760 gtacctctgcatggggccgcagcgaagatgcagggtgcttccagtccggaggaaga 819
QY 363 ctgcagcttcagaaactgctgtgagagcaggttacaatgtgtaccagtcgaagccca 422
DB 820 ctgtgctttcaggaggagatccgcccagatggctacaatgtgtaccgacccgagaaga 879
QY 423 tggcctgcccctgcgtctg 441
DB 880 ccgctcccggtctccctg 898

RESULT 10
AAC87022
ID AAC87022 standard; cDNA; 2137 BP.

XX AAC87022;

XX 20-APR-2001 (first entry)

XX Nucleotide sequence of human polypeptide PRO533.

XX Human; secreted protein; transmembrane protein; PRO196; PRO444; PRO183;
KW PRO185; PRO210; PRO215; PRO217; PRO242; PRO288; PRO365; PRO1361; PRO1308;
KW PRO1183; PRO1272; PRO1419; PRO4999; PRO7170; PRO248; PRO353; PRO1318;
KW PRO1600; PRO9940; PRO533; PRO301; PRO187; PRO337; PRO1411; PRO4356;
KW PRO246; PRO265; PRO941; PRO10096; PRO6003; PRO6004; PRO350; PRO2630;
KW PRO6309; cell death; genetic disorder; transgenic animal; gene therapy;
XX ss.

OS Homo sapiens.

XX Key Location/Qualifiers

PH CDS 464..1114

FT /*tag= a

FT sig_peptide 464..529
XX /*tag= b
XX WO200077037-A2.

XX 21-DEC-2000.

XX 22-MAY-2000; 2000WO-US14042.

XX 15-JUN-1999; 99US-0139695.

XX 20-JUL-1999; 99US-0145070.

XX 26-JUL-1999; 99US-0145698.

XX 17-AUG-1999; 99US-0149396.

XX 01-SEP-1999; 99WO-US20111.

XX 08-SEP-1999; 99WO-US20594.

XX 15-SEP-1999; 99WO-US21090.

XX 15-SEP-1999; 99WO-US21547.

XX 30-NOV-1999; 99WO-US28313.

XX 01-DEC-1999; 99WO-US28301.

XX 02-DEC-1999; 99WO-US28565.

XX 07-DEC-1999; 99US-0169495.

XX 05-JAN-2000; 2000WO-US00219.

XX 18-FEB-2000; 2000WO-US04341.

XX 18-FEB-2000; 2000WO-US04342.

XX 22-FEB-2000; 2000WO-US04414.

XX 01-MAR-2000; 2000WO-US05601.

XX 02-MAR-2000; 2000WO-US05841.

XX 20-MAR-2000; 2000WO-US07377.

XX 30-MAR-2000; 2000WO-US08439.

XX 15-MAY-2000; 2000WO-US13358.

XX 17-MAY-2000; 2000WO-US13705.

XX (GETH) GENENTECH INC.

XX Ashkenazi AJ, Baker KP, Botstein DA, Desnoyers L, Eaton DL;

XX Ferrara N, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;

XX Godowski PJ, Gurney AL, Kijavini IJ, Mather JP, Napier MA, Pan J;

XX Paoni NF, Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM;

XX Wood WI, Zhang Z;

XX WPI; 2001-050091/06.

XX P-PSDB; AAB31201.

XX Isolated nucleic acid molecule encoding a PRO polypeptide which is a
XX transmembrane polypeptide is useful for gene therapy and identification
XX of related polypeptides -

XX Claim 2; Fig 45; 244pp; English.

XX The present sequence encodes a human secreted and transmembrane
XX polypeptide. The specification describes human polypeptides, designated
XX PRO156, PRO444, PRO183, PRO185, PRO210, PRO215, PRO217, PRO242, PRO288,
XX PRO365, PRO1361, PRO1308, PRO1183, PRO1272, PRO1419, PRO4999, PRO7170,
XX PRO248, PRO353, PRO1318, PRO1600, PRO9940, PRO533, PRO301, PRO187,
XX PRO337, PRO1411, PRO4356, PRO246, PRO265, PRO941, PRO10096, PRO6003,
XX PRO6004, PRO350, PRO2630 and PRO6309. The biological activity of cells
XX can be modulated with agents that bind to these polypeptides, resulting
XX in the death of the cells. The polynucleotides encoding these
XX polypeptides are useful in the recombinant production of the
XX polypeptides, as a hybridisation probe to screen libraries to isolate
XX homologous sequences, or to map the gene. They may also be used for
XX analysing genetic disorders, and to produce transgenic animals which are
XX useful for the development and screening of therapeutically useful
XX reagents. The polynucleotides can also be used in gene therapy e.g. to
XX replace a defective gene.

XX Sequence 2137 BP; 422 A; 648 C; 598 G; 469 T; 0 other;

Query Match 8.4%; Score 54.2; DB 22; Length 2137;
Best Local Similarity 48.3%; Pred. No. 2e-05;
Matches 183; Conservative 0; Mismatches 193; Indels 3; Gaps 1;

KW Fibroblast growth factor-20; FGF-20; cytosolic; nontropic; metastasis;
 KW neuroprotective; antiparkinsonian; anticonvulsant; hepatotropic; cancer;
 KW leukaemia; tumour angiogenesis; skeletal dysplasia; Alzheimer's disease;
 KW neurodegenerative disorder; Parkinson's disease; multiple sclerosis;
 KW amyotrophic lateral sclerosis; progressive supranuclear palsy; epilepsy;
 KW Creutzfeldt-Jakob disease; hepatic disorder; cardiovascular disorder;
 KW haematopoietic disorder; myeloproliferative disorder; human; ss.
 XX
 OS Homo sapiens.
 XX
 XX
 PN WO200060085-A1.
 XX
 PD 12-OCT-2000.
 XX
 XX
 PF 27-MAR-2000; 2000WO-US08076.
 XX
 XX
 PR 02-APR-1999; 99US-0127534.
 PR 03-DEC-1999; 99US-0454470.
 XX
 XX (MILL-) MILLENNIUM PHARM INC.
 XX
 XX Curtis RAJ, Wrighton NC;
 XX
 DR WPI; 2000-647346/62.
 DR P-PSDB; AAY85639.
 XX
 XX Novel nucleic acids encoding fibroblast growth factor-20, useful for
 PT detecting and modulating FGF-20 and for treating, e.g. cancer, lar
 PT neurodegenerative disorders, hepatic disorders and cardiovascular
 PT disorders -
 XX
 XX Claim 1; Fig 10; 149pp; English.
 PS
 XX Polynucleotide sequences AAC61155-C61160 encode fibroblast growth
 CC factor-20 (FGF-20) proteins AAY85637-Y85639. The invention includes
 CC FGF-20 sequences isolated from humans and monkeys. Proteins,
 CC polynucleotides, agonists and antagonists of the invention have
 CC cytosolic; nontropic; neuroprotective; antiparkinsonian; anticonvulsant
 CC and hepatotropic activity. The peptides, nucleotides and their binding
 CC agents may be used to modulate the expression or activity of FGF-20, or to
 CC diagnose aberrant FGF-20 expression. FGF20 activity or expression may be
 CC down-regulated to treat proliferative or differentiative disorders (e.g.
 CC cancers and leukaemia), tumour angiogenesis and metastasis, skeletal
 CC dysplasia, neurodegenerative disorders (e.g. Alzheimer's disease,
 CC Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis,
 CC progressive supranuclear palsy, epilepsy or Creutzfeldt-Jakob disease),
 CC hepatic disorders, cardiovascular disorders and hematopoietic and/or
 CC myeloproliferative disorders.
 XX
 SQ Sequence 1973 BP; 580 A; 463 C; 505 G; 425 T; 0 other;

Query Match 7.4%; Score 48.2; DB 21; Length 1973;
 Best Local Similarity 52.2%; Pred. No. 0.00085;
 Matches 107; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
 QY 237 aagtcctcggagctcaagccttgaaagccaggggtcattcaaatcctgggtgtcaaacg 296
 DB 1061 aggtgcctcgtatcagatcagagagatgctggcttgggtgattacaggtgtgatgag 1120
 QY 297 ctctaggtttcttgcacacagcagatggagctctctatggatgcctcacttgatcc 356
 DB 1121 cagagatcactctcgtatcaggtatttcagagggcaacattttggatcacactatcgaccc 1180
 QY 357 tgagccctgcagctcagagaaactgctgtggagacggttacaatgtgtaccagctga 416
 DB 1181 ggagaactgcaggtttccacacacagacgctgtgaaacgggttaacgcttaccactctcc 1240
 QY 417 agcccatggcctgcctcgcgtctg 441
 DB 1241 teagtataccttccctggcagctcg 1265

RESULT 13
 AAC61157
 ID AAC61157 standard; cDNA; 2749 BP.
 XX
 AC AAC61157;
 XX
 DT 07-FEB-2001 (first entry)
 XX
 XX Human fibroblast growth factor-20 (FGF-20) cDNA sequence.

XX Fibroblast growth factor-20; FGF-20; cytosolic; nontropic; metastasis;
 KW neuroprotective; antiparkinsonian; anticonvulsant; hepatotropic; cancer;
 KW leukaemia; tumour angiogenesis; skeletal dysplasia; Alzheimer's disease;
 KW neurodegenerative disorder; Parkinson's disease; multiple sclerosis;
 KW amyotrophic lateral sclerosis; progressive supranuclear palsy; epilepsy;
 KW Creutzfeldt-Jakob disease; hepatic disorder; cardiovascular disorder;
 KW haematopoietic disorder; myeloproliferative disorder; human; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200060085-A1.
 XX
 PD 12-OCT-2000.
 XX
 XX
 PF 27-MAR-2000; 2000WO-US08076.
 XX
 XX
 PR 02-APR-1999; 99US-0127534.
 PR 03-DEC-1999; 99US-0454470.
 XX
 XX (MILL-) MILLENNIUM PHARM INC.
 XX
 XX Curtis RAJ, Wrighton NC;
 XX
 DR WPI; 2000-647346/62.
 DR P-PSDB; AAY85638.

XX Novel nucleic acids encoding fibroblast growth factor-20, useful for
 PT detecting and modulating FGF-20 and for treating, e.g. cancer,
 PT neurodegenerative disorders, hepatic disorders and cardiovascular
 PT disorders -
 XX
 XX Claim 1; Fig 9; 149pp; English.
 PS
 XX Polynucleotide sequences AAC61155-C61160 encode fibroblast growth
 CC factor-20 (FGF-20) proteins AAY85637-Y85639. The invention includes
 CC FGF-20 sequences isolated from humans and monkeys. Proteins,
 CC polynucleotides, agonists and antagonists of the invention have
 CC cytosolic; nontropic; neuroprotective; antiparkinsonian; anticonvulsant
 CC and hepatotropic activity. The peptides, nucleotides and their binding
 CC agents may be used to modulate the expression or activity of FGF-20, or to
 CC diagnose aberrant FGF-20 expression. FGF20 activity or expression may be
 CC down-regulated to treat proliferative or differentiative disorders (e.g.
 CC cancers and leukaemia), tumour angiogenesis and metastasis, skeletal
 CC dysplasia, neurodegenerative disorders (e.g. Alzheimer's disease,
 CC Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis,
 CC progressive supranuclear palsy, epilepsy or Creutzfeldt-Jakob disease),
 CC hepatic disorders, cardiovascular disorders and hematopoietic and/or
 CC myeloproliferative disorders.
 XX
 SQ Sequence 2749 BP; 747 A; 645 C; 633 G; 724 T; 0 other;

Query Match 7.4%; Score 48.2; DB 21; Length 2749;
 Best Local Similarity 52.2%; Pred. No. 0.00096;
 Matches 107; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
 QY 237 aagtcctcggagctcaagccttgaaagccaggggtcattcaaatcctgggtgtcaaacg 296
 DB 316 aggtgcctcgtatcagatcagagatgctggcttgggtgattacaggtgtgatgag 375
 QY 297 ctctaggtttcttgcacacagcagatggagctctctatggatgcctcacttgatcc 356
 DB 376 cagaagatcactcgtcgtgatttcagagggcaacattttggatcacactatcgaccc 435

CC dysplasia, neurodegenerative disorders (e.g. Alzheimer's disease,
CC Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis,
CC progressive supranuclear palsy, epilepsy or Creutzfeldt-Jakob disease),
CC hepatic disorders, cardiovascular disorders and haematopoietic and/or
CC myeloproliferative disorders.
XX

SQ Sequence 537 BP; 117 A; 173 C; 154 G; 93 T; 0 other;

```
Query Match          6.8%; Score 44.2; DB 21; Length 537;
Best Local Similarity 51.8%; Pred. No. 0.0069;
Matches 100; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 249 gctcaagccttgaagccagggtggtcattcaaatcctgggtgtcacaagcctctaggtttct 308
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3 gatcagatcagaggatgctggttgggtgattacaggtggtgatgagcagaataacct 62

QY 309 ttgccaacagccagatggagctctctatggatcgctcactttgatcctgagcctgcag 368
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 63 ctgcatggatttcagagggaacatttttggatcacactatttcgaccggagaaactgcag 122

QY 369 cttcagagaactgctgctggaggacggttacaatgtgtaccagttgaagcccatggcct 428
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 123 gttccaacaccagacgctggaacacgggtacgagctcaccactctcctcagtatcactt 182

QY 429 gccctgcgtctg 441
   | | | | | | | |
Db 183 cctggtcagtcg 195
```

Search completed: August 6, 2001, 22:30:58
Job time: 4904 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2001, 21:09:08 ; Search time 114.92 Seconds
(without alignments)
241.769 Million cell updates/sec

Title: US-09-391-861-4
Perfect score: 1115
Sequence: 1 MEWRSRVGTGLVRLLLA.....SSDPLSMVEPLQGRSPSYAS 210

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 133205027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPREMBL16:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1115	100.0	210	11	Q9JJN1	Q9jjn1 mus musculu
2	793.5	71.2	209	4	Q9NSAL	Q9nsal homo sapien
3	241	21.6	251	4	Q9GZV9	Q9gzv9 homo sapien
4	235	21.1	251	11	Q9EPC2	Q9epc2 mus musculu
5	212.5	19.1	227	13	Q9DDN0	Q9ddn0 gallus gall
6	173.5	15.6	191	13	Q9DFC9	Q9dfc9 brachydanio
7	171	15.3	206	13	Q9YGD8	Q9ygd8 oncorhynch
8	162	14.5	196	13	Q9YH31	Q9yh31 notophthalm
9	144	12.9	194	6	P79150	P79150 canis famil
10	144	12.9	212	11	Q9ESL9	Q9esl9 mus musculu
11	142.5	12.8	170	4	Q9HCT0	Q9hct0 homo sapien
12	141.5	12.7	212	13	O42407	O42407 gallus gall
13	139	12.5	162	11	Q9ESS2	Q9ess2 mus musculu
14	139	12.5	185	11	Q9ERN5	Q9ern5 rattus norv
15	139	12.5	211	4	Q9NP95	Q9np95 homo sapien
16	138	12.4	194	6	Q9N198	Q9n198 sus scrofa
17	138	12.4	212	11	Q9EST9	Q9est9 rattus norv
18	135.5	12.2	213	6	Q9N1B9	Q9n1b9 ovis aries
19	135	12.1	207	11	Q9ESL8	Q9esl8 mus musculu

20	135	12.1	207	11	Q9ERQ5	Q9erq5 mus musculu
21	133	11.9	130	6	O77767	O77767 canis famil
22	132	11.8	196	4	P78443	P78443 homo sapien
23	131	11.7	208	13	Q9PVY1	Q9pvy1 xenopus lae
24	129	11.6	170	11	O60487	O60487 cavia porce
25	125	11.2	425	5	O76831	O76831 caenorhabdi
26	124	11.1	199	13	Q9IAI3	Q9iai3 gallus gall
27	123	11.0	245	13	Q9W6A2	Q9w6a2 gallus gall
28	123	11.0	252	11	O89096	O89096 rattus norv
29	121	10.9	192	11	Q9ERW3	Q9erw3 rattus norv
30	121	10.9	237	13	Q9IAI6	Q9iai6 gallus gall
31	121	10.9	253	13	Q9IAI5	Q9iai5 gallus gall
32	118.5	10.6	108	6	Q9N1S7	Q9n1s7 capreolus c
33	118	10.6	106	6	Q9N1S8	Q9n1s8 capreolus c
34	116	10.4	101	13	P79706	P79706 cynops pyrr
35	115	10.3	127	4	Q99517	Q99517 homo sapien
36	115	10.3	163	11	Q9JHL9	Q9jhl9 mus musculu
37	113.5	10.2	146	13	O07659	O07659 gallus gall
38	113.5	10.2	181	13	Q9IAI7	Q9iai7 gallus gall
39	113.5	10.2	243	13	Q9W6A1	Q9w6a1 gallus gall
40	110	9.9	97	4	Q9NSJ0	Q9nsj0 homo sapien
41	110	9.9	192	4	O95830	O95830 homo sapien
42	109.5	9.8	64	11	Q9JK55	Q9jk55 rattus norv
43	99.5	8.9	76	6	Q9NOV2	Q9nov2 ovis aries
44	98.5	8.8	74	6	O77561	O77561 oryctolagus
45	98	8.8	770	5	P91672	P91672 drosophila

ALIGNMENTS

RESULT 1
Q9JJN1
ID Q9JJN1 PRELIMINARY; PRT: 210 AA.
AC Q9JJN1;
DT 01-OCR-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE FGF-21.
GN FGF-21.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20461777; PubMed-10858549;
RA Nishimura T., Nakatake Y., Konishi M., Itoh N.;
RT "Identification of a novel FGF, FGF-21, preferentially expressed in
RT the liver.";
RL Biochim. Biophys. Acta 1492:203-206(2000).
DR EMBL; AB025718; BAA99416.1; -;
DR InterPro; IPR002209; -;
DR InterPro; IPR002348; -;
DR Pfam; PF00167; FGF; 1.
DR PRINTS; PR00263; HBGFFGF.
DR PRINTS; PR00262; IL1HBGF.
DR SMART; SM00442; FGF; 1.
SQ SEQUENCE 210 AA; 23237 MW; AE02ABA6477E6F0 CRC64;

Query Match 100.0%; Score 1115; DB 11; Length 210;
Best Local Similarity 100.0%; Pred. No. 8.7e-96;
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEWRSRVGTGLVRLLLAVFLGVQTAIPIDSSPLQFGGVQVQRVRYLTDQDTEA 60
Db 1 MEWRSRVGTGLVRLLLAVFLGVQTAIPIDSSPLQFGGVQVQRVRYLTDQDTEA 60
QY 61 HLEIREDTGVVGAHRSPESLLELKALKPGVITLGVKASRFLCCQPDGALYGSHPDPE 120
Db 61 HLEIREDTGVVGAHRSPESLLELKALKPGVITLGVKASRFLCCQPDGALYGSHPDPE 120

```
QY 121 ACSFRELLEDGYNVYQSAHGLPLRLPKQDSPNODATSWGVPRLPMPGLLHEPDQAG 180
Db 121 ACSFRELLEDGYNVYQSAHGLPLRLPKQDSPNODATSWGVPRLPMPGLLHEPDQAG 180
QY 181 FLPPPPDVGSSDPLSMVPELQGRSPSYAS 210
Db 181 FLPPPPDVGSSDPLSMVPELQGRSPSYAS 210

RESULT 2
Q9NSAI PRELIMINARY; PRT; 209 AA.
AC Q9NSAI;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE FGF-21.
GN FGF-21.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20461777; PubMed=10858549;
RA Nishimura T., Nakatake Y., Konishi M., Itoh N.;
RT "Identification of a novel FGF, FGF-21, preferentially expressed in
the liver.";
RL Biochim. Biophys. Acta 1492:203-206(2000).
DR EMBL; AB021975; BAA99415.1; -.
DR InterPro; IPR002209; -.
DR Pfam; PF00167; FGF; 1.
DR PRINTS; PR00263; HBGF; FGF.
DR SMART; SM00442; FGF; 1.
SQ SEQUENCE 209 AA; 22300 MW; 27925C52A0023823 CRC64;
```

```
Query Match 71.2%; Score 793.5; DB 4; Length 209;
Best Local Similarity 78.9%; Pred. No. 5.8e-66;
Matches 157; Conservative 10; Mismatches 31; Indels 1; Gaps 1;
```

```
QY 12 GLWVRLLLAVFLGLGVYQAVPIPDSSPLLOFGGVQRORYLYTDDQDTEAHLEIREGTVV 71
Db 12 GLWVS-VLGLLGLACQAHPIPDSSPLLOFGGVQRORYLYTDDQDTEAHLEIREGTVG 70
QY 72 GAAHRSPELLELKALKPGVIQILGVKASRFLCQDPGALYGSHPDFPEACSFRELLLED 131
Db 71 GAADQSPESLQLKALKPGVIQILGVKTSRFLCQDPGALYGSHPDFPEACSFRELLLED 130
QY 132 GYNVYQSEAHGLPLRLPKQDSPNODATSWGVPRLPMPGLLHEPDQAGFLPPEPDVGS 191
Db 131 GYNVYQSEAHGLPLRLPKQDSPNODATSWGVPRLPMPGLLHEPDQAGFLPPEPDVGS 190
QY 192 SDPLSMVPELQGRSPSYAS 210
Db 191 SDPLSMVPELQGRSPSYAS 209
```

```
RESULT 3
Q9GZV9 PRELIMINARY; PRT; 251 AA.
AC Q9GZV9;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE FGF-23 (FGF23).
GN FGF23.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11032749;
RA Yamashita T., Yoshioka M., Itoh N.;
RT "Identification of a novel fibroblast growth factor, FGF-23,
preferentially expressed in the ventrolateral thalamic nucleus of the
brain.";
RL Biochem. Biophys. Res. Commun. 277:494-498(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20517346; PubMed=11062477;
RA White K.E., Evans W.E., O'Riordan J.L.H., Speer M.C., Econs M.J.,
RA Lorenz-Depiereux B., Grabowski M., Meitinger T., Strom T.M.;
RT "Autosomal dominant hypophosphataemic rickets is associated with
mutations in FGF23.";
RL Nat. Genet. 26:345-348(2000).
DR EMBL; AB037973; BAB13477.1; -.
DR EMBL; AF263537; AAG09917.1; -.
SQ SEQUENCE 251 AA; 27954 MW; 6093BD0CC50C2489 CRC64;
```

```
Query Match 21.6%; Score 241; DB 4; Length 251;
Best Local Similarity 32.3%; Pred. No. 1.3e-14;
Matches 71; Conservative 38; Mismatches 71; Indels 40; Gaps 8;
```

```
QY 11 LGLWVRLLLAVFLGLGVYQAVPIPDSSPLLOFGGVQRORYLYTDDQDTEAHLEIREG 68
Db 6 LRLWVLCALCSVCMSVLRAY--PNASPLGSSWGGLI---HLTATARN--YHLQIHKG 59
QY 69 TVVGAHRSPELLELKALKPGVIQILGVKASRFLCQDPGALYGSHPDFPEACSFRELL 128
Db 60 HVDGAPHQTIYSALMIRSEDAGFVVTGVMSRRYLCMDFRGNIFGSHYFDPENCRFQHOT 119
QY 129 LEDGYNVYQS-EAHGLPLRLPKQDSPNODATSWGVPRLPMPGLLHEPDQAG----- 178
Db 120 LENGVDVYHSPQYHFL-----VSLGRAKRAFLPGMNPYPYQFLSRNEIP 165
QY 179 -AGFLPPEP-----PDVGSDDLPMVPELQGRSPSYAS 210
Db 166 LIHFWTPPRHRSERDSDSERDPLNLKPRARMTAPAS 205
```

```
RESULT 4
Q9EPC2 PRELIMINARY; PRT; 251 AA.
AC Q9EPC2;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE FGF-23 (FGF23).
GN FGF23.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11032749;
RA Yamashita T., Yoshioka M., Itoh N.;
RT "Identification of a novel fibroblast growth factor, FGF-23,
preferentially expressed in the ventrolateral thalamic nucleus of the
brain.";
RL Biochem. Biophys. Res. Commun. 277:494-498(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE-WHOLE EMBRYO;
RX MEDLINE=20517346; PubMed=11062477;
RA White K.E., Evans W.E., O'Riordan J.L.H., Speer M.C., Econs M.J.,
RA Lorenz-Depiereux B., Grabowski M., Meitinger T., Strom T.M.;
RT "Autosomal dominant hypophosphataemic rickets is associated with
mutations in FGF23.";
RL Nat. Genet. 26:345-348(2000).
DR EMBL; AB037889; BAB13478.1; -.

```


[illegible]

```
DR EMBL; AB049218; BAB16406.1; -.
SQ SEQUENCE 212 AA; 23659 MW; 174DBCE8913B69EF CRC64;

Query Match 12.9%; Score 144; DB 11; Length 212;
Best Local Similarity 35.3%; Pred. No. 1.1e-05;
Matches 53; Conservative 12; Mismatches 65; Indels 20; Gaps 7;

QY 33 PDSSPLQFGGQVRQRYLYTDDQDTEAHLEIREDTGVVGAHRSPEL---LELKALP 89
DB 52 FGSVELAHHLGILRRRLQYLC-----RTGPHLQILPDGTVGQGT--RQDHSFGILEFISVAV 105
QY 90 GVIQTLGVKASRFLCQPDGALYGSPPHDFPEACSFRELLLEDGYNVYQSE--AHGLPLR- 146
DB 106 GLVSRGVDGLYLGMDKGLYSEKLTSE-CIFREQFEENWYNTYSSNIYKHGNTGRR 164
QY 147 ----LPQKDSPNQDATSWGP---VRFLPMP 169
DB 165 YFVALNKDGTDRDGARSRRQKFTHFLRP 194

RESULT 11
Q9HCT0 PRELIMINARY; PRT; 170 AA.
AC Q9HCT0;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE FGF-22.
GN FGF22.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Itoh N.;
RT "human FGF-22.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB01925; BAB13479.1; -.
SQ SEQUENCE 170 AA; 19662 MW; CB88918C2D54ACE7 CRC64;

Query Match 12.8%; Score 142.5; DB 4; Length 170;
Best Local Similarity 33.1%; Pred. No. 1.1e-05;
Matches 52; Conservative 24; Mismatches 66; Indels 15; Gaps 8;

QY 4 MRSRVGTGLG-WVRLLAFLVLLGVYQAYPIPDSSPLLQFGGQVRQRYLYTDDQDTEAHL 62
DB 1 MRRL-WLGLAWLLARAPDAAGTSPASRGRSYPHLE--GDVWRRLFS-----STHFFL 53
QY 63 EIREDTGVVGAHR-SPESLLELKALKPGVIOILGVKASRFLCQPDGALYGSPPHDFPEA 121
DB 54 RVDPPGVRGQGTWRHGGQDSILRSVGVVVIKAVSSGFFYVAMNRRGLYGSRLYTVD- 112
QY 122 CSPRELLLEDGYNVYQSE---AHGLP--LRLPQKDS 153
DB 113 CRFRERIEENGHTYASQRWRRRGQPMFLALDRGGP 149

RESULT 12
O42407 PRELIMINARY; PRT; 212 AA.
AC O42407;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE FTBROBLAST GROWTH FACTOR 10.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OX NCBI_TaxID=9031;

RN SEQUENCE FROM N.A.
RP MEDLINE=97330690; PubMed=9187149;
RA Ouchi H., Nakagawa T., Yamamoto A., Araga A., Ohta T., Ishimaru Y.,
RA Yoshio H., Kuwana T., Nohno T., Yanasaki M., Itoh N., Noji S.;
RA "The mesenchymal factor, FGF10, initiates and maintains the outgrowth
RT of the chick limb bud through interaction with FGF8, an apical
RT ectodermal factor.";
RL Development 124:2235-2244(1997).
DR EMBL; D86333; BAA24945.1; -.
DR HSSP; P03968; IBAR.
DR InterPro; IPR002209; -.
DR PFam; PF00167; FGF; 1.
DR PRINTS; PR00263; HBGFFGF.
DR ProDom; PD000831; -.
DR SMART; SM00442; FGF; 1.
SQ SEQUENCE 212 AA; 23631 MW; AB4C0B32C72A0D90 CRC64;

Query Match 12.7%; Score 141.5; DB 13; Length 212;
Best Local Similarity 31.1%; Pred. No. 1.8e-05;
Matches 38; Conservative 27; Mismatches 38; Indels 19; Gaps 5;

QY 29 AYPIDSSPL-----LQFGGQVRQRYLYTDDQDTEAHLEIREDTGVVGAHRS 77
DB 56 SFPSSPSSPSAGRHVRSYNHLQ--GDVRRKLYSYN---KYFLKIEKNGKVGSTKEN 109
QY 78 -PESLLELKALKPGVIOILGVKASRFLCQPDGALYGSPPHDFPEACSFRELLLEDGYNVY 136
DB 110 CPTSLIEITSVEIGVAVKSIKSNYILAMNKGKVGSKFNSD-CKLKERIEENYNY 168
QY 137 QS 138
DB 169 AS 170

RESULT 13
Q9ESS2 PRELIMINARY; PRT; 162 AA.
AC Q9ESS2;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE FGF-22.
GN FGF-22.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Itoh N.;
RT "mouse FGF-22.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB036765; BAB16407.1; -.
SQ SEQUENCE 162 AA; 18927 MW; 225EF512F4E1BE29 CRC64;

Query Match 12.5%; Score 139; DB 11; Length 162;
Best Local Similarity 30.9%; Pred. No. 2.2e-05;
Matches 55; Conservative 30; Mismatches 55; Indels 38; Gaps 11;

QY 4 MRSRVGTGLG-WVRLLAFLVLLGVYQAYPIPDSSPLLQFGGQVRQRYLYTDDQDTEAHL 62
DB 1 MRSRL-WLGLAW--LLLA-----RAPGAPGGYPHLE--GDVWRRLF-----SSTHF 42
QY 63 EIREDTGVVGAHR-SPESLLELKALKPGVIOILGVKASRFLCQPDGALYGSPPHDF 119
DB 43 FLRVDLGRVQGTWRHGGQDSIVRSVGVVVIKAVSGFFYVAMNRRGLYGSRYSV 102
QY 120 EACSFRELLLEDGYNVYQSE-----EAHGLPLRLPKQDSPNODATSWGPV 163
DB 103 D-CRFRERIEENGHTYASRWRHRRGPMFLALDSQGIIP-RQGRTRRHQSTHFLPV 158
```

```
DR EMBL; AB049218; BAB16406.1; -.
SQ SEQUENCE 212 AA; 23659 MW; 174DBCE8913B69EF CRC64;

Query Match 12.9%; Score 144; DB 11; Length 212;
Best Local Similarity 35.3%; Pred. No. 1.1e-05;
Matches 53; Conservative 12; Mismatches 65; Indels 20; Gaps 7;

QY 33 PDSSPLQFGGQVRQRYLYTDDQDTEAHLEIREDTGVVGAHRSPEL---LELKALP 89
DB 52 FGSVELAHHLGILRRRLQYLC-----RTGPHLQILPDGTVGQGT--RQDHSFGILEFISVAV 105
QY 90 GVIQTLGVKASRFLCQPDGALYGSPPHDFPEACSFRELLLEDGYNVYQSE--AHGLPLR- 146
DB 106 GLVSRGVDGLYLGMDKGLYSEKLTSE-CIFREQFEENWYNTYSSNIYKHGNTGRR 164
QY 147 ----LPQKDSPNQDATSWGP---VRFLPMP 169
DB 165 YFVALNKDGTDRDGARSRRQKFTHFLRP 194

RESULT 11
Q9HCT0 PRELIMINARY; PRT; 170 AA.
AC Q9HCT0;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE FGF-22.
GN FGF22.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Itoh N.;
RT "human FGF-22.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB01925; BAB13479.1; -.
SQ SEQUENCE 170 AA; 19662 MW; CB88918C2D54ACE7 CRC64;

Query Match 12.8%; Score 142.5; DB 4; Length 170;
Best Local Similarity 33.1%; Pred. No. 1.1e-05;
Matches 52; Conservative 24; Mismatches 66; Indels 15; Gaps 8;

QY 4 MRSRVGTGLG-WVRLLAFLVLLGVYQAYPIPDSSPLLQFGGQVRQRYLYTDDQDTEAHL 62
DB 1 MRRL-WLGLAWLLARAPDAAGTSPASRGRSYPHLE--GDVWRRLFS-----STHFFL 53
QY 63 EIREDTGVVGAHR-SPESLLELKALKPGVIOILGVKASRFLCQPDGALYGSPPHDFPEA 121
DB 54 RVDPPGVRGQGTWRHGGQDSILRSVGVVVIKAVSSGFFYVAMNRRGLYGSRLYTVD- 112
QY 122 CSPRELLLEDGYNVYQSE---AHGLP--LRLPQKDS 153
DB 113 CRFRERIEENGHTYASQRWRRRGQPMFLALDRGGP 149

RESULT 12
O42407 PRELIMINARY; PRT; 212 AA.
AC O42407;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE FTBROBLAST GROWTH FACTOR 10.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OX NCBI_TaxID=9031;
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2001, 21:09:57 ; Search time 36.96 Seconds

(without alignments)
194.633 Million cell updates/sec

Title: US-09-391-861-4

Perfect score: 1115

Sequence: 1 MEMWRSRVGTGLWVRLLA.....SSDPLSMVEPLOGRSPSYAS 210

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwlseqProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	262.5	23.5	216	1	FGFJ_HUMAN
2	207.5	18.6	218	1	FGFJ_MOUSE
3	178.5	16.0	208	1	FGF6_MOUSE
4	177.5	15.9	208	1	FGF6_HUMAN
5	171.5	15.4	206	1	FGF4_BOVIN
6	170.5	15.3	239	1	FGF3_HUMAN
7	170.0	15.2	187	1	FGF4_XENLA
8	169.5	15.2	245	1	FGF3_MOUSE
9	166.5	14.9	194	1	FGF4_CHICK
10	165.5	14.8	206	1	FGF4_HUMAN
11	160.5	14.4	192	1	FGF4_XENLA
12	158.5	14.2	256	1	FGF3_BRARE
13	155.5	13.9	202	1	FGF4_MOUSE
14	149.5	13.4	237	1	FGF3_XENLA
15	149.5	13.4	264	1	FGF5_MOUSE
16	145.5	13.0	220	1	FGF3_CHICK
17	145.0	13.0	194	1	FGF7_SHEEP
18	145.0	13.0	266	1	FGF5_RAT
19	143.0	12.8	194	1	FGF7_HUMAN
20	142.5	12.8	155	1	FGF1_MESAU
21	142.0	12.7	194	1	FGF7_MOUSE
22	142.0	12.7	208	1	FGF4_HUMAN
23	142.0	12.7	215	1	FGF4_RAT
24	141.5	12.7	155	1	FGF1_MOUSE
25	140.0	12.6	209	1	FGF4_MOUSE
26	140.0	12.6	268	1	FGF5_HUMAN
27	135.0	12.1	207	1	FGF3_HUMAN
28	134.0	12.0	154	1	FGF2_MOUSE
29	134.0	12.0	154	1	FGF2_RAT
30	133.5	12.0	155	1	FGF1_HUMAN
31	133.0	11.9	155	1	FGF2_BOVIN
32	133.0	11.9	155	1	FGF2_SHEEP
33	133.0	11.9	194	1	FGF7_RAT

34	132.5	11.9	155	1	FGF1_BOVIN	P03968	bos taurus
35	132.0	11.8	155	1	FGF2_HUMAN	P09038	homo sapien
36	131.5	11.8	208	1	FGF9_MOUSE	P31371	homo sapien
37	131.5	11.8	208	1	FGF9_MOUSE	P54130	mus musculus
38	131.5	11.8	208	1	FGF9_RAT	P36364	rattus norv
39	131.0	11.7	156	1	FGF2_MONDO	P48798	monodelphis
40	131.0	11.7	207	1	FGFG_RAT	O54769	rattus norv
41	131.0	11.7	209	1	FGF9_XENLA	Q91875	xenopus lae
42	129.5	11.6	152	1	FGF1_PIG	P20002	sus scrofa
43	128.0	11.5	155	1	FGF2_XENLA	P12226	xenopus lae
44	126.0	11.3	158	1	FGF2_CHICK	P48800	gallus gall
45	125.0	11.2	413	1	L7556_CAEEL	Q11184	caenorhabdi

ALIGNMENTS

```

RESULT 1
FGFJ_HUMAN
ID FGFJ_HUMAN STANDARD; PRT; 216 AA.
AC O95750;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE FIBROBLAST GROWTH FACTOR-19 PRECURSOR (FGF-19).
GN FGF19.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
EX MEDLINE=99132028; PubMed=9931477;
RA Nishimura T., Utsunomiya Y., Hoshikawa M., Ohuchi H., Itoh N.:
RT "Structure and expression of a novel human FGF, FGF-19, expressed in
RT the fetal brain.";
RL Biochim. Biophys. Acta 1444:148-151(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Xie M.-H., Holcomb I., Deuel B., Dowd P., Huang A., Vagts A.,
RA Foster J., Brush J., Gu Q., Liang J., Hillan K., Goddard A.,
RA Gurney A.L.;
RT "FGF-19, a novel fibroblast growth factor with unique specificity for
RT FGF4.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -! FUNCTION: MAY BE INVOLVED IN BRAIN DEVELOPMENT DURING
CC EMBRYOGENESIS.
CC -! SUBCELLULAR LOCATION: SECRETED.
CC -! TISSUE SPECIFICITY: EXPRESSED ONLY IN FETAL BRAIN.
CC -! SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB018122; BAA75500.1; -
CC EMBL; AF110400; AAD45973.1; -
CC MIM; 603891; -
CC HSSP; P09038; 2BPH.
CC InterPro; IPR002209; -
CC InterPro; IPR002348; -
CC Pfam; PF00167; FGF; 1.
CC PRINTS; PR00262; ILIHGFG.
CC PRINTS; PR00263; HBGFGFG.
CC PROSITE; PS00247; HBGFG_FGF; 1.
CC Growth factor; Signal.
CC SIGNAL 1 22
CC CHAIN 23 216
CC POTENTIAL.
CC FIBROBLAST GROWTH FACTOR-19.

```

4


```
DR PRINTS; PR00263; HBGFGF.
DR PROSITE; PS00247; HBGFGF; 1.
KW Transforming protein; Oncogene; Growth factor; Mitogen; Signal;
KW Glycoprotein.
FT SIGNAL 1 ? POTENTIAL.
FT CHAIN 1 ? 208 FIBROBLAST GROWTH FACTOR-6.
FT CARBOHYD 45 45 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFID 90 157 POTENTIAL.
SQ SEQUENCE 208 AA; 22798 MW; BB7B9301BB79A1B CRC64;

Query Match 16.0%; Score 178.5; DB 1; Length 208;
Best Local Similarity 35.1%; Pred. No. 8.5e-09;
Matches 47; Conservative 20; Mismatches 50; Indels 17; Gaps 5;

QY 43 GQVRQRYLYTDDQDTEAHLEIREDTGVVGAHRSPELLEKALKPGVQIILGVKASRF 102
DB 81 GIKRQRRLYC--NVGIGFHLQVPPDGRISGTHEENPYSLEISTVERGVSLFGVKSALF 138
QY 103 LCQQPDGALYGSFHPDPEACSFRELLLEDGYNVYOSEAH-GLPLRLPQ-----KDSP 153
DB 139 IANNSKGRLYTTPSFIDE-CKFREILLPNPNYAYESDLYRGYIALSKYGRVKRGSKVSP 197
QY 154 NQDATSWGPFVRLP 167
DB 198 IMTVT-----HFLP 206

RESULT 4
FGF6_HUMAN STANDARD; PRT; 208 AA.
ID FGF6_HUMAN
AC P10767;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE FIBROBLAST GROWTH FACTOR-6 PRECURSOR (FGF-6) (HBGF-6) (HST-2).
GN FGF6 OR HST2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91360279; PubMed=1886714;
RA Coullier F., Batoz M., Marics I., de Lapeyriere O., Birnbaum D.;
RT "Putative structure of the FGF6 gene product and role of the signal
RT peptide.";
RL Oncogene 6:1437-1444 (1991).
RN [2]
RP SEQUENCE OF 11-208 FROM N.A.
RX MEDLINE=92195660; PubMed=1549352;
RA Lida S., Yoshida T., Naito K., Sakamoto H., Katoh O., Hirohashi S.,
RA Sato T., Onda M., Sugimura T., Terada M.;
RT "Human hst-2 (FGF-6) oncogene: cDNA cloning and characterization.";
RL Oncogene 7:303-309 (1992).
RN [3]
RP SEQUENCE OF 81-208 FROM N.A.
RX MEDLINE=89201880; PubMed=2649847;
RA Marics I., Adelaide J., Raybaud F., Mattei M.-G., Coullier F.,
RA Planche J., de Lapeyriere O., Birnbaum D.;
RT "Characterization of the HST-related FGF.6 gene, a new member of the
RT fibroblast growth factor gene family.";
RL Oncogene 4:335-340 (1989).
CC -1- FUNCTION: CAN TRANSFORM NIH 3T3 CELLS. EXHIBITS STRONG MITOGENIC
CC AND ANGIOGENIC PROPERTIES.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- TISSUE SPECIFICITY: LEUKEMIA CELL LINES WITH PLATELET/
CC MEGAKARYOCYTIC DIFFERENTIATION POTENTIAL.
CC -1- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; X57075; CAA40358.1;
CC EMBL; X57075; CAA40359.1; ALT_INIT.
CC EMBL; X57075; CAA40360.1; ALT_INIT.
CC EMBL; X63454; CAA45054.1;
CC EMBL; X14071; CAB37648.2;
CC EMBL; X14072; CAB37648.2; JOINED.
CC EMBL; X14073; CAB37648.2; JOINED.
CC PIR; S04204; S04204.
CC PIR; S23739; S23739.
CC HSP; P09038; IBLD.
CC MIM; 134921;
CC InterPro; IPR002209;
CC InterPro; IPR002348;
CC Pfam; PF00167; FGF; 1.
CC PRINTS; PR00262; IL1HBGF.
CC PRINTS; PR00263; HBGFGF.
CC PROSITE; PS00247; HBGFGF; 1.
KW Transforming protein; Oncogene; Growth factor; Mitogen; Signal;
KW Glycoprotein.
FT SIGNAL 1 ? POTENTIAL.
FT CHAIN 1 ? 208 FIBROBLAST GROWTH FACTOR-6.
FT CARBOHYD 45 45 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFID 90 157 POTENTIAL.
FT CONFLICT 100 100 V->G (IN REF. 3).
SQ SEQUENCE 208 AA; 22904 MW; 79EF44685B324322 CRC64;

Query Match 15.9%; Score 177.5; DB 1; Length 208;
Best Local Similarity 34.3%; Pred. No. 1e-08;
Matches 46; Conservative 21; Mismatches 50; Indels 17; Gaps 5;

QY 43 GQVRQRYLYTDDQDTEAHLEIREDTGVVGAHRSPELLEKALKPGVQIILGVKASRF 102
DB 81 GIKRQRRLYC--NVGIGFHLQVPPDGRISGTHEENPYSLEISTVERGVSLFGVKSALF 138
QY 103 LCQQPDGALYGSFHPDPEACSFRELLLEDGYNVYOSEAH-GLPLRLPQ-----KDSP 153
DB 139 VAMNSKGRLYATPSFQ-ECKFERETLLPNPNYAYESDLYQGYIALSKYGRVKRGSKVSP 197
QY 154 NQDATSWGPFVRLP 167
DB 198 IMTVT-----HFLP 206

RESULT 5
FGF4_BOVIN STANDARD; PRT; 206 AA.
ID FGF4_BOVIN
AC P48803;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE FIBROBLAST GROWTH FACTOR-4 PRECURSOR (FGF-4) (HEPARIN SECRETORY
DE TRANSFORMING PROTEIN) (HST) (HBGF-4).
GN FGF4 OR HST.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96032369; PubMed=7557455;
RA Yu J.C., Dessebra A.J., Wang L.M., Fleming T.P., Chedid M.,
RA Miki T., Heidar M.A.;
RT "An unexpected transforming gene in calf-thymus carrier DNA: bovine
RT hst.";
RL Gene 162:333-334 (1995).
```


DR HSP; P09038; 2BFH.
 DR InterPro; IPR002209; -.
 DR InterPro; IPR002348; -.
 DR Pfam; PF00167; FGF; 1.
 DR PRINTS; PR00262; IL1HBGF.
 DR PRINTS; PR00263; HBGF-FGF.
 DR PROSITE; PS00247; HBGF_FGF; 1.
 KW Growth factor; Mitogen; Signal.
 FT SIGNAL 1 ? POTENTIAL.
 FT CHAIN ? 187 FIBROBLAST GROWTH FACTOR-4-I.
 SQ SEQUENCE 187 AA; 21223 MW; AA63D65E82AD1BD CRC64; 6;

Query Match 15.2%; Score 170; DB 1; Length 187;
 Best Local Similarity 34.3%; Pred. No. 4.1e-08;
 Matches 49; Conservative 22; Mismatches 52; Indels 20; Gaps 6;

QY 34 DSSPLQFGQVQRVLYTDDQDTEAHLEIREDTGVVGAHRSPELLELKALPGVIQ 93
 DB 54 DSDYLL---GIRQRRLYC--NVGIGFHIQVLPDGRINGMHSNRYSLLELSPVEVGVS 108
 QY 94 ILGVKASRLCQPDGALYGSDFEDPEACSFRELLEDGYNVYQSEAH-GLPLRLPQ--- 149
 DB 109 LYGVKSGMFVAMNAKGLYSRYFN-ECKFKETLLPNNNAYESRKYPGMYIALGKNR 167

QY 150 -----KDSPNQDATSGPVRFLP 167
 DB 168 TTKGNVSPWTILT-----HFLLP 185

RESULT 8
 FGF3_MOUSE STANDARD; PRT; 245 AA.
 AC P05524;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE INT-2 PROTO-ONCOGENE PROTEIN PRECURSOR (FIBROBLAST GROWTH FACTOR-3)
 DE (FGF-3) (HBGF-3).
 GN FGF3 OR FGF-3 OR INT-2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86247582; PubMed=3013624;
 RA Moore R., Casey G., Brookes S., Dixon M., Peters G., Dickson C.;
 RT "Sequence, topography and protein coding potential of mouse int-2: a putative oncogene activated by mouse mammary tumour virus."; EMBO J. 5:919-924(1986).
 RL [2]
 RN SEQUENCE OF 1-73 FROM N.A.
 RX MEDLINE=88296404; PubMed=2841106;
 RA Smith R., Peters G., Dickson C.;
 RT "Multiple RNAs expressed from the int-2 gene in mouse embryonal carcinoma cell lines encode a protein with homology to fibroblast growth factors."; EMBO J. 7:1013-1022(1988).
 RL [3]
 RN CHARACTERIZATION.
 RP MEDLINE=91193291; PubMed=1964688;
 RX Dickson C., Acland P., Smith R., Dixon M., Deed R., McAllan D., Walther W., Fuller-Pace F., Kiefer P., Peters G.;
 RT "Characterization of int-2: a member of the fibroblast growth factor family."; J. Cell Sci. Suppl. 13:87-96(1990).
 RL J. Cell Sci. Suppl. 13:87-96(1990).
 CC -!- FUNCTION: COULD BE INVOLVED IN EAR DEVELOPMENT.
 CC -!- INDUCTION: BY INTEGRATION OF MOUSE MAMMARY TUMOR VIRUS.
 CC -!- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -----

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>) or send an email to license@isb-sib.ch.

 DR EMBL; Y00848; CAA68767.1; -.
 DR PIR; A23930; TVMST2.
 DR HSP; P09038; 2BFH.
 DR MGD; MGI:95517; Fgf3.
 DR InterPro; IPR002209; -.
 DR InterPro; IPR002348; -.
 DR Pfam; PF00167; FGF; 1.
 DR PRINTS; PR00262; IL1HBGF.
 DR PRINTS; PR00263; HBGF-FGF.
 DR PROSITE; PS00247; HBGF_FGF; 1.
 KW Transforming protein; Oncogene; Growth factor; Mitogen; Signal;
 KW Glycoprotein.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 245 INT-2 PROTO-ONCOGENE PROTEIN.
 FT CARBOHYD 65 65 N-LINKED (GLCNAC...).
 SQ SEQUENCE 245 AA; 27214 MW; 70D94FD6A7837C79 CRC64;

Query Match 15.2%; Score 169.5; DB 1; Length 245;
 Best Local Similarity 27.5%; Pred. No. 6.4e-08;
 Matches 61; Conservative 21; Mismatches 63; Indels 77; Gaps 8;

QY 25 GYQAYPIPDSSPLQFGQVQRVLYTDDQDTEAHLEIREDTGVVGAHRSPELLEL 84
 DB 34 GVEY-----HLGGAPRRRLYC-----ATKYHLQLHPGSRVNGSLNSAYSILEI 78
 QY 85 KALPGVIQILGVKASRLCQPDGALYGSDFEDPEACSFRELLEDGYNVYQSEAH--- 141
 DB 79 TAVEGVVAINGLFSGRVLAAMKRGRLYADHYNAE-CEFVERIHELGYNTIAGRLRTG 137
 QY 142 -----GLPLRLPQKDSNQDATSGPVRFLP----- 167
 DB 138 SSGPCAQRPQCAQRPWYVSVNGKGRPRGFKTRTKSS-----LFLPVLGHK 186

QY 168 ---MPGLLHEPDQA---GFLPPE-----PPDVGSSDPLS 196
 DB 187 DHEVMVRLQSSQAPRAGSGSQPRQRORRKKQSPGDHGRMETLS 228

RESULT 9
 FGF4_CHICK STANDARD; PRT; 194 AA.
 AC P48804;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE FIBROBLAST GROWTH FACTOR-4 PRECURSOR (FGF-4) (HBGF-4).
 GN FGF4 OR FGF-4.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95021713; PubMed=7935794;
 RA Niswander L., Jeffrey S., Martin G.R., Tickle C.;
 RT "A positive feedback loop coordinates growth and patterning in the vertebrate limb."; Nature 371:609-612(1994).
 RL Nature 371:609-612(1994).
 CC -!- FUNCTION: MAY BE THE ENDOGENOUS RIDGE SIGNAL FOR LIMB DEVELOPMENT.
 CC RECIPROCAL INTERACTIONS MAY CREATE A POSITIVE FEEDBACK LOOP BETWEEN SONIC HEDGEHOG (SHH) AND FGF4.
 CC -!- TISSUE SPECIFICITY: POSTERIOR RIDGE.
 CC -!- INDUCTION: BY RETINOIC ACID.
 CC -!- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; U14654; AAA58706.1; -.
 DR HSP; P09038; 1BFF.
 DR InterPro; IPR002209; -.
 DR InterPro; IPR002348; -.
 DR Pfam; PF00167; FGF; 1.
 DR PRINTS; PR00262; ILIHGFG.
 DR PRINTS; PR00263; HBGFFGFG.
 DR PROSITE; PS00247; HBGF_FGF; 1.
 KW Growth factor; Mitogen; Signal.
 FT SIGNAL 1 ? POTENTIAL.
 FT CHAIN 1 ? 194 FIBROBLAST GROWTH FACTOR-4.
 SQ SEQUENCE 194 AA; 21642 MW; 5D44245BE97E664A CRC64;

Query Match 14.9%; Score 166.5; DB 1; Length 194;
 Best Local Similarity 32.8%; Pred. No. 8.8e-08;
 Matches 44; Conservative 22; Mismatches 51; Indels 17; Gaps 5;

QY 43 GQVRYLYTDDQDQDAHLEIREDTGVVGAHRSPELLEKALKPGVIOILGVKASRF 102

DB 67 GYKRLRLYC--NVGIFHIQVLPGRDIGHSENRYSLLEISPVGVSVFGRSLGF 124

QY 103 LCOQPDGALYSPHPDPEACSFRELLLEDGYNVYQSEAH-GLPLRLPQ-----KDSP 153

DB 125 VAMNSKGKLYGTHVND-CKEKEILLPNNNAYESRIYPGMYIALSKNGRTKGNKVSP 183

QY 154 NODATSWGVPVRLP 167

DB 184 TMTVT-----HFLP 192

RESULT 10

FGF4_HUMAN

ID FGF4_HUMAN STANDARD; PRT; 206 AA.
 AC P08620;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE FIBROBLAST GROWTH FACTOR-4 PRECURSOR (FGF-4) (HEPARIN SECRETORY
 DE TRANSFORMING PROTEIN) (HST-1) (HST) (TRANSFORMING PROTEIN KS3) (HBGF-
 DE 4).
 GN FGF4 OR HST OR HSTF1 OR KS3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

SEQUENCE FROM N.A.

RX MEDLINE=88041096; PubMed=2959959;
 RA Yoshida T., Miyagawa K., Odagiri H., Sakamoto H., Little P.F.R.,
 RA Terada M., Sugimura T.;
 RT Genomic sequence of hst, a transforming gene encoding a protein
 RT homologous to fibroblast growth factors and the int-2-encoded
 RT protein.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:7305-7309(1987).
 RN [2]

SEQUENCE FROM N.A.

RX MEDLINE=87204251; PubMed=2953031;
 RA Taira M., Yoshida T., Miyagawa K., Sakamoto H., Terada M.,
 RA Sugimura T.;
 RT "cDNA sequence of human transforming gene hst and identification of
 RT the coding sequence required for transforming activity.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:2980-2984(1987).
 RN [3]

SEQUENCE FROM N.A.

RX MEDLINE=87301716; PubMed=2957062;
 RA Delli Bovi P., Curatola A.M., Kern F.G., Greco A., Ittmann M.,
 RA Basilico C.;
 RT "An oncogene isolated by transfection of Kaposi's sarcoma DNA encodes
 RT a growth factor that is a member of the FGF family.";
 RL Cell 50:729-737(1987).
 CC -1- FUNCTION: CAN TRANSFORM NIH 3T3 CELLS FROM A HUMAN STOMACH TUMOR
 CC (HST) AND FROM KAPOSI'S SARCOMA (KS3). IT HAS A MITOGENIC
 CC ACTIVITY.

CC -1- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; J02986; AAB59555.1; -.
 DR EMBL; M17446; AAA59473.1; -.
 DR PIR; A28417; TVHUHS.
 DR HSP; P09038; 1BFF.
 DR MIM; 164980; -.
 DR InterPro; IPR002209; -.
 DR InterPro; IPR002348; -.
 DR Pfam; PF00167; FGF; 1.
 DR PRINTS; PR00262; ILIHGFG.
 DR PRINTS; PR00263; HBGFFGFG.
 DR PROSITE; PS00247; HBGF_FGF; 1.
 KW Transforming protein; Oncogene; Growth factor; Mitogen; Signal.
 FT SIGNAL 1 ? POTENTIAL.
 FT CHAIN 1 ? 206 FIBROBLAST GROWTH FACTOR-4.
 SQ SEQUENCE 206 AA; 22047 MW; C7FD54A0272A1569 CRC64;

Query Match 14.8%; Score 165.5; DB 1; Length 206;
 Best Local Similarity 36.2%; Pred. No. 1.2e-07;
 Matches 42; Conservative 19; Mismatches 40; Indels 15; Gaps 4;

QY 61 HLEIREDTGVVGAHRSPELLEKALKPGVIOILGVKASRFLLCQPDGALYSPHPDPE 120

DB 95 HLQALPDGRIGGAHADTRDLSLSPVGVSVIFGVASRFVAMSKGLYSGPFTDE 154

QY 121 ACSFRELLLEDGYNVYQSEAH-GLPLRLPQ-----KDSPNODATSWGVPVRLP 167

DB 155 -CFKEILLPNNNAYESKYPCGFALSKNGTKGNRVSPMTKVT-----HFLP 204

RESULT 11

FGFB_XENLA

ID FGFB_XENLA STANDARD; PRT; 192 AA.
 AC P48806;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE FIBROBLAST GROWTH FACTOR-4-II PRECURSOR (FGF-4-II) (HBGF-4-II)
 DE (EMBRYONIC FIBROBLAST GROWTH FACTOR II) (XERGF-II).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]

SEQUENCE FROM N.A.

RX MEDLINE=92315916; PubMed=1618138;
 RA Isaacs H.V., Tannahill D., Slack J.M.W.;
 RT "Expression of a novel FGF in the Xenopus embryo. A new candidate
 RT inducing factor for mesoderm formation and anteroposterior
 RT specification.";
 RL Development 114:711-720(1992).

CC -1- FUNCTION: GOOD CANDIDATE FOR AN INDUCING FACTOR WITH POSSIBLE
 CC ROLES BOTH IN MESODERM INDUCTION AT THE BLASTULA STAGE AND IN THE


```

DR HSP; P09038; 2BFH.
DR MGD; MG1:95518; Fgf4.
DR InterPro; IPR002209; -.
DR PRINTS; PR002348; -.
DR Pfam; PF00167; FGF; 1.
DR PRINTS; PR00262; IL1HBGF.
DR PRINTS; PR00263; HBGF_FGF.
DR PROSITE; PS00247; HBGF_FGF; 1.
DR TRANSFORMING PROTEIN; Oncogene; Growth factor; Mitogen; Signal.
FT SIGNAL 1 ? POTENTIAL.
FT CHAIN 1 ? FIBROBLAST GROWTH FACTOR-4.
FT CONFLICT 167 167 A -> S (IN REF. 2).
FT SEQUENCE 202 AA; 62D456231047CA31 CRC64;

Query Match 13.9%; Score 155.5; DB 1; Length 202;
Best Local Similarity 33.6%; Pred. No. 8.5e-07;
Matches 39; Conservative 21; Mismatches 41; Indels 15; Gaps 4;

QY 61 HLEIREDTGVVGAHRSPELLEKALKPGVILGVKASRFLCQPDGALYGSFHDPE 120
|::: || | : : ||||| : : || : | : : | : | : | : | : | : | : |
Db 91 HLQVLPDGRIGGVHADTDSLLELSPVQGVVSIFGVASRFFVMSRSGKLGVPFFETDE 150
|::: || | : : ||||| : : || : | : : | : | : | : | : | : | : |
QY 121 ACSFRELLEDGYNVQSEAH-GLPLRLPQ-----KDSPNQDATSGPVRFLP 167
|::: || | : : ||||| : : || : | : : | : | : | : | : | : | : |
Db 151 -CKEKEILLPNNYAYEAYPGFMALSKNGRTKKGNRVSPMTKVT-----HFLP 200

RESULT 14
FGF3_XENLA
ID FGF3_XENLA STANDARD; PRT; 237 AA.
AC P36386;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE FIBROBLAST GROWTH FACTOR-3 PRECURSOR (FGF-3) (HBGF-3) (INT-2).
GN FGF3.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=94038898; PubMed=8223431;
RA Kiefer P., Mathieu M., Close J.M., Peters G., Dickson C.;
RT "FGF3 from Xenopus laevis.";
RL EMBO J. 12:4159-4168(1993).
RN [2]
RP SEQUENCE OF 39-137 FROM N.A.
RC TISSUE=Neurula;
RX MEDLINE=93048831; PubMed=1425349;
RA Tannahill D., Isaacs H.V., Close M.J., Peters G., Slack J.M.W.;
RT "Developmental expression of the Xenopus int-2 (FGF-3) gene:
activation by mesodermal and neural induction.";
RL Development 115:695-702(1992).
CC -!- FUNCTION: POTENT MITOGEN AND TRANSFORMING AGENT.
CC -!- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z25539; CAA80987.1; -.
CC EMBL; X65237; CAA46341.1; -.
CC PIR; S39582; S39582.
CC PIR; S25713; S25713.
CC HSP; P09038; 2BFH.
CC InterPro; IPR002209; -.

```

```

DR InterPro; IPR002348; -.
DR Pfam; PF00167; FGF; 1.
DR PRINTS; PR00262; IL1HBGF.
DR PRINTS; PR00263; HBGF_FGF.
DR PROSITE; PS00247; HBGF_FGF; 1.
DR GROWTH factor; Mitogen; Signal; Glycoprotein.
FT SIGNAL 1 21
FT CHAIN 22 237 FIBROBLAST GROWTH FACTOR-3.
FT CARBOHYD 83 83 N-LINKED (GLCNAC. ).
FT SEQUENCE 237 AA; 26984 MW; EDD31B08933567A2D CRC64;

Query Match 13.4%; Score 149.5; DB 1; Length 237;
Best Local Similarity 33.9%; Pred. No. 3.4e-06;
Matches 40; Conservative 18; Mismatches 43; Indels 17; Gaps 4;

QY 25 GYQAYPIPDSPLLQFGQVQRVRYLYTDDQDTEAHLEIREDTGVVGAHR-SPELLE 83
|::: || | : : ||||| : : || : | : : | : | : | : | : | : | : |
Db 52 GYVE-----HLGGAPNRKLYC-----ATKYHLQIHLNGKINGTLEKNSVFSILE 96
|::: || | : : ||||| : : || : | : : | : | : | : | : | : | : |
QY 84 LKALKPGVILGVKASRFLCQPDGALYGSFHDPEACSFRELLEDGYNVQSEAH 141
|::: || | : : ||||| : : || : | : : | : | : | : | : | : | : |
Db 97 ITAVDVGIVAIKGLFSGRYLAMNQRGLYASETYNPE-CEFEVRIHELGYNTYASRLY 153
|::: || | : : ||||| : : || : | : : | : | : | : | : | : | : |

RESULT 15
FGF5_MOUSE
ID FGF5_MOUSE STANDARD; PRT; 264 AA.
AC P15656;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE FIBROBLAST GROWTH FACTOR-5 PRECURSOR (FGF-5) (HBGF-5).
GN FGF5 OR FGF-5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90201563; PubMed=2318343;
RA Hebert J.M., Basilio C., Goldfarb M., Haub O., Martin G.R.;
RT "Isolation of cDNAs encoding four mouse FGF family members and
characterization of their expression patterns during embryogenesis.";
RL Dev. Biol. 138:454-463(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=91045929; PubMed=1700424;
RA Haub O., Drucker B., Goldfarb M.;
RT "Expression of the murine fibroblast growth factor 5 gene in the
adult central nervous system.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8022-8026(1990).
CC -!- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M30643; AAA96698.1; -.
CC EMBL; M37823; AAB02660.1; -.
CC EMBL; M37821; AAB02660.1; JOINED.
CC EMBL; M37822; AAB02660.1; JOINED.
CC EMBL; M37821; AAB02659.1; ALT_SEQ.
CC PIR; A36207; A36207.
CC PIR; B37360; B37360.
CC HSP; P09038; IBFC.
CC MGD; MG1:95519; Fgf5.
CC InterPro; IPR002209; -.

```

```
DR InterPro: IPR002348; -.
DR Pfam: PF00167; FGF; 1.
DR PRINTS; PR00262; ILIHGCF.
DR PRINTS; PR00263; HBGFEGF.
DR PROSITE; PS00247; HBGF_FGF; 1.
KW Transforming protein; Oncogene; Growth factor; Mitogen; Signal.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 264 FIBROBLAST GROWTH FACTOR-5.
FT DOMAIN 53 59 POLY-SER.
FT CARBOHYD 108 108 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 264 AA; 29103 MW; F6A9C8153EE923D1 CRC64;

Query Match 13.4%; Score 149; DB 1; Length 264;
Best Local Similarity 25.0%; Pred. No. 4.3e-06;
Matches 52; Conservative 26; Mismatches 80; Indels 50; Gaps 6;

QY 35 SSPLQFGGQVRQRYLYTDDQDEA-----HLEIREDTGVVGAHRSPESLLELK 85
Db 63 ASFGSGSGSEHSFOWSPSGRTGSLYCRVGIGFHLQIYPDGKVGNSHEASVLSILEIF 122
QY 86 ALKPGVIQILGVKASRFLCQOPDGLYGSFHPDPEACSFRELLEDGYNVYOSEAH---- 141
Db 123 AVSQGIVGIRGVFSNKFAMSKKGLHASAKFTDD-CKFRERFQENSNTYIYASAIHRTK 181
QY 142 -----GLPLRLPQKDSNQDATSWGVPVRFLLPMPGLLHEPQDQAGFL-- 182
Db 182 TGREWYVALMKRGKAKRGCSRV----KQHVSTHF-----LPRFKQSEQPELSFTVT 230
QY 183 -----PPEPPDYGSSDPLSMVEPLOGR 204
Db 231 VPEKKAPVPKVPKPLSQPRRSPSPVKYR 258
```

Search completed: August 6, 2001, 21:09:58
Job time: 724 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
OM protein - protein search, using sw model
Run on: August 6, 2001, 21:07:01 ; Search time 67.44 Seconds
(without alignments)
237.198 Million cell updates/sec
Title: US-09-391-861-4
Perfect score: 1115
Sequence: 1 MEWMRSRVGTGLWVRLLA.....SSDPLSMVEPLQGRSPSYAS 210
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 219241
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : PIR_68: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	235	21.1 251 3 JC7513	fibroblast growth
2	178.5	16.0 208 2 S14192	fibroblast growth
3	177.5	15.9 208 2 S20102	fibroblast growth
4	170.5	15.3 239 1 S04742	fibroblast growth
5	170	15.2 187 2 S23595	embryonic fibroblast growth
6	169.5	15.2 245 1 TWAST2	transforming prote
7	167.5	15.0 206 2 JC4268	fibroblast growth
8	166.5	14.9 194 2 I50710	fibroblast growth
9	165.5	14.8 206 1 TVHUHS	fibroblast growth
10	160.5	14.4 192 2 S34407	embryonic fibroblast growth
11	158.5	14.2 256 2 JC4627	fibroblast growth
12	155.5	13.9 202 1 TWASHS	fibroblast growth
13	149.5	13.4 237 1 S39582	transforming prote
14	149	13.4 264 2 S36207	fibroblast growth
15	147	13.2 194 2 S26049	fibroblast growth
16	145.5	13.0 220 2 I50588	fibroblast growth
17	145	13.0 194 2 S49501	keratinocyte growth
18	145	13.0 266 2 S68144	fibroblast growth
19	143	12.8 194 1 A36301	fibroblast growth
20	142.5	12.8 155 1 A60721	acidic fibroblast growth
21	142	12.7 194 2 I48610	keratinocyte growth
22	141.5	12.7 155 2 S04147	acidic fibroblast growth
23	141.5	12.7 155 2 D37360	acidic fibroblast growth
24	139.5	12.5 267 1 TVHUES	fibroblast growth
25	139	12.5 211 2 JC7353	fibroblast growth
26	138	12.4 212 3 JC7511	fibroblast growth
27	135	12.1 207 2 JC5941	fibroblast growth
28	134	12.0 154 2 A31674	basic fibroblast growth
29	134	12.0 154 2 C37360	basic fibroblast growth

30	133.5	12.0	155	1	A33665	acidic fibroblast
31	133	11.9	146	1	S00185	basic fibroblast g
32	133	11.9	157	1	GKBOB	basic fibroblast g
33	132.5	11.9	155	1	GKBOB	acidic fibroblast
34	132	11.8	210	2	A32398	basic fibroblast g
35	131.5	11.8	208	2	S66486	fibroblast growth
36	131.5	11.8	208	2	A48137	fibroblast growth
37	131	11.7	164	2	S31622	basic fibroblast g
38	131	11.7	207	2	JC5940	fibroblast growth
39	131	11.7	208	2	JH0782	fibroblast somatot
40	129.5	11.6	152	2	JH0476	acidic fibroblast
41	128	11.5	155	1	A40117	basic fibroblast g
42	126	11.3	189	2	A48834	basic fibroblast g
43	124.5	11.2	155	2	A60130	acidic fibroblast
44	123.5	11.1	137	2	I46711	fibroblast growth
45	123.5	11.1	155	2	JW0055	acidic fibroblast

ALIGNMENTS

RESULT 1
JC7513
fibroblast growth factor-23 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-2001 #sequence_revision 31-Mar-2001 #text_change 31-Mar-2001
C:Accession: JC7513
R:Yamashita, T.; Yoshioka, M.; Itoh, N.
Biochem. Biophys. Res. Commun. 277, 494-498, 2000
A:Title: Identification of a novel fibroblast growth factor, FGF-23, preferentially e
A:Reference number: JC7513; MUID:20490027
A:Contents: Embryo
A:Accession: JC7513
A:Molecule type: mRNA
A:Residues: 1-251 <YAM>
A:Cross-references: DDBJ:AB037889
C:Comment: This factor, a secreted protein, acts on proximal cells and plays roles in
C:Genetics:
A:Gene: fgf-23
C:Keywords: nucleus

Query Match 21.1%; Score 235; DB 3; Length 251;
Best Local Similarity 34.4%; Pred. No. 1.2e-13;
Matches 73; Conservative 30; Mismatches 81; Indels 28; Gaps 9;
QY 11 LGLWVRLLAFLVFL---LGVYQAYIPDSSPLL--QFGGVQRORYLYTDDQDTEAHLEI 64
Db 2 LGTCLRLLVGLCTVCSLGTARAY--PDTSPLLGNNGLSLT---HLYTATAR-TSYHLQI 55
QY 65 REDGTVVGAHRSPESSLELKALKPGVQIILGVKASRFLCQPDGALYSGSPHDPACSF 124
Db 56 HRDHVGTGTPHTYTSALMTSDAGSVITGATRRFLCMDLHGNIFGSLHSPENCKF 115
QY 125 RELLEDGYNVYQSEAHGLPLRLPKDQSPNQATSWGPV-RFLPMPG---LLH-----E 174
Db 116 RQWLENGYDVLVSKHLYLVSLGRKRIFQPGNPPFPQFLARRNEVLLHFFYTVRPR 175
QY 175 PQDQAGFLPEPPVGVGSDPLSMVEPLQGRSP 206
Db 176 RHTRSADPPPE-----RDPLNLVKPRPRATP 201

RESULT 2
S14192
fibroblast growth factor 6 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 17-Mar-2000
R:de Lapeyriere, O.; Rosnet, O.; Benharroch, D.; Raybaud, F.; Marchetto, S.; Planche,
Oncogene 5, 823-831, 1990
A:Title: Structure, chromosome mapping and expression of the murine Fgf-6 gene.
A:Reference number: S14192; MUID:90295275

QY 84 LKALRPGV IQLGLVAKSR LCCQYDFGALIGSFHDFEACSEABDDUDDENIVYQSEHH 141

97 ITAYDVGVAIKGLFSGRYLANNEKGRLYASEVFNRE-CEFLERIHELGYNTYASRHH 153

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2001, 21:05:42 ; Search time 18.95 Seconds
(without alignments)
228.178 Million cell updates/sec

Title: US-09-391-861-4
Perfect score: 1115
Sequence: 1 MEMRSRVGTGLWVRLLA.....SSDPLSMVEPQGRSPSYAS 210

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2.6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2.6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2.6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2.6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2.6/prodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2.6/prodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	180	16.1	158	3	US-08-776-207-18
2	177.5	15.9	145	1	US-08-551-171-3
3	177.5	15.9	145	1	US-08-902-233-3
4	177.5	15.9	146	1	US-08-551-171-4
5	177.5	15.9	146	3	US-08-902-233-4
6	177.5	15.9	170	1	US-08-551-171-1
7	177.5	15.9	170	3	US-08-902-233-1
8	177.5	15.9	171	1	US-08-551-171-2
9	177.5	15.9	171	3	US-08-902-233-2
10	177.5	15.9	174	1	US-08-551-171-7
11	177.5	15.9	174	3	US-08-902-233-7
12	177.5	15.9	175	1	US-08-551-171-8
13	177.5	15.9	175	3	US-08-902-233-8
14	177.5	15.9	198	1	US-08-439-725A-14
15	177.5	15.9	198	1	US-08-441-629-12
16	177.5	15.9	198	2	US-08-867-471-14
17	177.5	15.9	198	2	US-08-438-439C-10
18	177.5	15.9	198	2	US-08-438-439C-20
19	177.5	15.9	198	3	US-08-718-904-15
20	177.5	15.9	198	3	US-09-023-082A-14
21	177.5	15.9	198	3	US-08-776-207-12
22	177.5	15.9	198	5	PCT-US95-09172-12
23	177.5	15.9	207	1	US-08-551-171-5
24	177.5	15.9	207	3	US-08-902-233-5
25	177.5	15.9	208	1	US-08-551-171-6
26	177.5	15.9	208	1	US-08-464-590A-16
27	177.5	15.9	208	1	US-08-462-169B-14

28	177.5	15.9	208	2	US-08-207-412B-11	Sequence 11, Appl
29	177.5	15.9	208	2	US-08-951-822-27	Sequence 27, Appl
30	177.5	15.9	208	3	US-09-103-079-14	Sequence 14, Appl
31	177.5	15.9	208	3	US-08-902-233-6	Sequence 6, Appl
32	177.5	15.9	208	3	US-09-093-585-16	Sequence 16, Appl
33	170.5	15.3	239	1	US-08-464-590A-11	Sequence 11, Appl
34	170.5	15.3	239	1	US-08-462-169B-11	Sequence 11, Appl
35	170.5	15.3	239	2	US-08-207-412B-14	Sequence 14, Appl
36	170.5	15.3	239	2	US-08-951-822-35	Sequence 35, Appl
37	170.5	15.3	239	3	US-03-103-079-11	Sequence 11, Appl
38	170.5	15.3	239	3	US-08-718-904-12	Sequence 12, Appl
39	170.5	15.3	239	3	US-09-023-082A-21	Sequence 21, Appl
40	170.5	15.3	239	3	US-09-093-585-11	Sequence 11, Appl
41	170	15.2	187	1	US-08-441-629-11	Sequence 11, Appl
42	170	15.2	187	3	US-08-776-207-11	Sequence 11, Appl
43	170	15.2	187	5	PCT-US95-09172-11	Sequence 11, Appl
44	169.5	15.2	245	1	US-08-439-725A-11	Sequence 11, Appl
45	169.5	15.2	245	2	US-08-867-471-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-08-776-207-18
; Sequence 18, Application US/08776207A
; Patent No. 6080718
; GENERAL INFORMATION:
; APPLICANT: Kirschner, Marc W.
; APPLICANT: Kinoshita, No. 6080718iyuki
; TITLE OF INVENTION: Receptor-Ligand Assay
; FILE REFERENCE: HU95-01A2
; CURRENT APPLICATION NUMBER: US/08/776.207A
; CURRENT FILING DATE: 1997-06-23
; EARLIER APPLICATION NUMBER: PCT/US95/09172
; EARLIER FILING DATE: 1995-07-19
; EARLIER APPLICATION NUMBER: 08/441.629
; EARLIER FILING DATE: 1995-05-15
; EARLIER APPLICATION NUMBER: 08/279.217
; EARLIER FILING DATE: 1994-07-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial
US-08-776-207-18

Query Match	16.1%	Score 180:	DB 3:	Length 158;
Best Local Similarity	39.3%	Pred. No. 1.le-11;		
Matches	53;	Conservative	12;	Mismatches 50; Indels 20; Gaps 5;
QY	43	GOVQRYLITDDQDTEAHLEITREDGTGVGAHRSPELSLELKALPKGVIOILGVKASRF	102	
Db	33	GKKRTRRLYC---RVGGFHLQILPDGRVNGTHESNRYLSLELSAVEGVWSIKGVESGLF	89	
QY	103	LCQQPDGALYSGPHDPACSFRELLLEDGYNVYQS-----EAHGLPLRLPKQDS	152	
Db	90	LAMNKKGLYASKKP-TTECKFKERLENNYNTYASAKRYGMYVALNKNRPKR-GSKTS	147	
QY	153	PNQDATSGNRPVRFPLP	167	
Db	148	PTQKAT-----HFLP	157	

RESULT 2
US-08-551-171-3
; Sequence 3, Application US/08551171
; Patent No. 5679550
; GENERAL INFORMATION:

```
; APPLICANT: YOSHIMURA, Koji
; APPLICANT: ISHIMARU, Kaori
; APPLICANT: IGARASHI, Koichi
; APPLICANT: TERADA, Masaaki
; TITLE OF INVENTION: hst-2 MUTEIN, DNA CODING FOR THE SAME AND
; TITLE OF INVENTION: PREPARATION THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/551.171
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/196.001
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: NEUNER, George W
; REGISTRATION NUMBER: 26964
; REFERENCE/DOCKET NUMBER: 44146
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)523-3400
; TELEFAX: (617)523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-551-171-3

Query Match 15.9%; Score 177.5; DB 1; Length 145;
Best Local Similarity 34.3%; Pred. No. 1.8e-11;
Matches 46; Conservative 21; Mismatches 50; Indels 17; Gaps 5;

QY 43 GQVRQRYLYTDDQDTEAHLEIREDTGVVGAHRSPESLLELKALPGVQIILGVKASRF 102
| ||| || : ||: || : ||: || : ||: || : ||: || : ||: || : ||
Db 18 GIKRQRRLYC--NVGIGFHLQVLPDGRISGTHEENPYSLLLEISTVERGVSLFGVRSALF 75

QY 103 LCQQPDGALYSGPHDPACSFRELLLEDGYNVYQSEAH-GLPLRLPQ-----KDSP 153
: ||| || : ||: || : ||: || : ||: || : ||: || : ||: || : ||
Db 76 VAMNSKGRLYATPSFQ-ECKFRETLLPNNNAYESDLYOGTYIALSKYGRVKRGSKVSP 134

QY 154 NQDATSWGPFVRFLP 167
| |||
Db 135 IMTVT-----HFLP 143

RESULT 3
US-08-902-233-3
; Sequence 3, Application US/08902233
; Patent No. 6013784
; GENERAL INFORMATION:
; APPLICANT: YOSHIMURA, Koji
; APPLICANT: ISHIMARU, Kaori
; APPLICANT: IGARASHI, Koichi
; APPLICANT: TERADA, Masaaki
; TITLE OF INVENTION: hst-2 MUTEIN, DNA CODING FOR THE SAME AND
; TITLE OF INVENTION: PREPARATION THEREOF
; NUMBER OF SEQUENCES: 10
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/902.233
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/196.001
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: NEUNER, George W
; REGISTRATION NUMBER: 26964
; REFERENCE/DOCKET NUMBER: 44146
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)523-3400
; TELEFAX: (617)523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-902-233-3

Query Match 15.9%; Score 177.5; DB 3; Length 145;
Best Local Similarity 34.3%; Pred. No. 1.8e-11;
Matches 46; Conservative 21; Mismatches 50; Indels 17; Gaps 5;

QY 43 GQVRQRYLYTDDQDTEAHLEIREDTGVVGAHRSPESLLELKALPGVQIILGVKASRF 102
| ||| || : ||: || : ||: || : ||: || : ||: || : ||: || : ||
Db 18 GIKRQRRLYC--NVGIGFHLQVLPDGRISGTHEENPYSLLLEISTVERGVSLFGVRSALF 75

QY 103 LCQQPDGALYSGPHDPACSFRELLLEDGYNVYQSEAH-GLPLRLPQ-----KDSP 153
: ||| || : ||: || : ||: || : ||: || : ||: || : ||: || : ||
Db 76 VAMNSKGRLYATPSFQ-ECKFRETLLPNNNAYESDLYOGTYIALSKYGRVKRGSKVSP 134

QY 154 NQDATSWGPFVRFLP 167
| |||
Db 135 IMTVT-----HFLP 143

RESULT 4
US-08-551-171-4
; Sequence 4, Application US/08551171
; Patent No. 5679550
; GENERAL INFORMATION:
; APPLICANT: YOSHIMURA, Koji
; APPLICANT: ISHIMARU, Kaori
; APPLICANT: IGARASHI, Koichi
; APPLICANT: TERADA, Masaaki
; TITLE OF INVENTION: hst-2 MUTEIN, DNA CODING FOR THE SAME AND
; TITLE OF INVENTION: PREPARATION THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
```



```
;
;
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/551,171
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/196,001
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: NEUNER, George W
; REGISTRATION NUMBER: 26964
; REFERENCE/DOCKET NUMBER: 44146
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)523-3400
; TELEFAX: (617)523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 146 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-551-171-4.

Query Match 15.9%; Score 177.5; DB 1; Length 146;
Best Local Similarity 34.3%; Pred. No. 1.8e-11;
Matches 46; Conservative 21; Mismatches 50; Indels 17; Gaps 5;

QY 43 GOVROLYTDDQDTEAHLEIREDTGVVGAHRSPELLELKALPGVILGVKASRF 102
Db 19 GKRRRLYC--NVGIGFHLQVLPDGRISGTHEENPYSLEISTVERGVVSLFGVRSALF 76
QY 103 LCQQPDGALYGFHPDPCAGSFRELLLEDGYNVYQSEAH-GLPLRLPQ-----KDSP 153
Db 77 VAMNSKGLRYATPSFQ-EECKFERETLLPNNYNAYESDLYQGTIYALSXYGRVKGSKVSP 135
QY 154 NQDATSWGVPVRLP 167
Db 136 IMVT-----HFLP 144

RESULT 5
US-08-902-233-4
; Sequence 4, Application US/08902233
; Patent No. 6013784
; GENERAL INFORMATION:
; APPLICANT: YOSHIMURA, Koji
; APPLICANT: ISHIMARU, Kaori
; APPLICANT: IGARASHI, Koichi
; APPLICANT: TERADA, Masaaki
; TITLE OF INVENTION: hst-2 MUTEIN, DNA CODING FOR THE SAME AND
; PREPARATION THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

```
;
;
; APPLICATION NUMBER: US/08/902,233
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/196,001
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: NEUNER, George W
; REGISTRATION NUMBER: 26964
; REFERENCE/DOCKET NUMBER: 44146
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)523-3400
; TELEFAX: (617)523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 146 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-902-233-4

Query Match 15.9%; Score 177.5; DB 3; Length 146;
Best Local Similarity 34.3%; Pred. No. 1.8e-11;
Matches 46; Conservative 21; Mismatches 50; Indels 17; Gaps 5;

QY 43 GOVROLYTDDQDTEAHLEIREDTGVVGAHRSPELLELKALPGVILGVKASRF 102
Db 19 GKRRRLYC--NVGIGFHLQVLPDGRISGTHEENPYSLEISTVERGVVSLFGVRSALF 76
QY 103 LCQQPDGALYGFHPDPCAGSFRELLLEDGYNVYQSEAH-GLPLRLPQ-----KDSP 153
Db 77 VAMNSKGLRYATPSFQ-EECKFERETLLPNNYNAYESDLYQGTIYALSXYGRVKGSKVSP 135
QY 154 NQDATSWGVPVRLP 167
Db 136 IMVT-----HFLP 144

RESULT 6
US-08-551-171-1
; Sequence 1, Application US/08551171
; Patent No. 5679550
; GENERAL INFORMATION:
; APPLICANT: YOSHIMURA, Koji
; APPLICANT: ISHIMARU, Kaori
; APPLICANT: IGARASHI, Koichi
; APPLICANT: TERADA, Masaaki
; TITLE OF INVENTION: hst-2 MUTEIN, DNA CODING FOR THE SAME AND
; PREPARATION THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/551,171
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/196,001
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
```

```

; NAME: NEUNER, George W
; REGISTRATION NUMBER: 26964
; REFERENCE/DOCKET NUMBER: 44146
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)523-3400
; TELEFAX: (617)523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 170 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-551-171-1

Query Match 15.9%; Score 177.5; DB 1; Length 170;
Best Local Similarity 34.3%; Pred. No. 2.3e-11;
Matches 46; Conservative 21; Mismatches 50; Indels 17; Gaps 5;

QY 43 GOVQRRLYTDQDTEAHLEIREDTGVVGAHRAHSPESLLELKALKPGVIOILGVKASRF 102
| ||| || : ||: || : ||: || : ||: || : ||: || : ||: || : ||: ||
Db 43 GIKRRLVC--NVGIGFHLQVLPDGRISGTHEENPYSILLESTVERGVSLFGVRSALF 100
: ||| || : ||: || : ||: || : ||: || : ||: || : ||: || : ||: ||
QY 103 LCOOPDGLYSGPHFDEACSFRELLEDGYNVYQSEAH-GLPLRLPQ-----KDSP 153
: ||| || : ||: || : ||: || : ||: || : ||: || : ||: || : ||: ||
Db 101 VAMNSKGRLYATPSFQ-EECKFRETLPPNNYAYESDLYQGTYIALSKYGRVKRGSKVSP 159
| ||| || : ||: || : ||: || : ||: || : ||: || : ||: || : ||: ||
QY 154 NQDATSWGVPVRLP 167
| | | | |
Db 160 IMTVT-----HELP 168

RESULT 7
US-08-902-233-1
; Sequence 1, Application US/08902233
; Patent No. 6013784
; GENERAL INFORMATION:
; APPLICANT: YOSHIMURA, Koji
; APPLICANT: ISHIMARU, Kaori
; APPLICANT: IGARASHI, Koichi
; APPLICANT: TERADA, Masaaki
; TITLE OF INVENTION: hst-2 MUTEIN, DNA CODING FOR THE SAME AND
; PREPARATION THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/902.233
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/196.001
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: NEUNER, George W
; REGISTRATION NUMBER: 26964
; REFERENCE/DOCKET NUMBER: 44146
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)523-3400
; TELEFAX: (617)523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 171 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-551-171-2

Query Match 15.9%; Score 177.5; DB 3; Length 170;
Best Local Similarity 34.3%; Pred. No. 2.3e-11;
Matches 46; Conservative 21; Mismatches 50; Indels 17; Gaps 5;

QY 43 GOVQRRLYTDQDTEAHLEIREDTGVVGAHRAHSPESLLELKALKPGVIOILGVKASRF 102
| ||| || : ||: || : ||: || : ||: || : ||: || : ||: || : ||: ||
Db 43 GIKRRLVC--NVGIGFHLQVLPDGRISGTHEENPYSILLESTVERGVSLFGVRSALF 100
: ||| || : ||: || : ||: || : ||: || : ||: || : ||: || : ||: ||
QY 103 LCOOPDGLYSGPHFDEACSFRELLEDGYNVYQSEAH-GLPLRLPQ-----KDSP 153
: ||| || : ||: || : ||: || : ||: || : ||: || : ||: || : ||: ||
Db 101 VAMNSKGRLYATPSFQ-EECKFRETLPPNNYAYESDLYQGTYIALSKYGRVKRGSKVSP 159
| ||| || : ||: || : ||: || : ||: || : ||: || : ||: || : ||: ||
QY 154 NQDATSWGVPVRLP 167
| | | | |
Db 160 IMTVT-----HELP 168

RESULT 8
US-08-551-171-2
; Sequence 2, Application US/08551171
; Patent No. 5679550
; GENERAL INFORMATION:
; APPLICANT: YOSHIMURA, Koji
; APPLICANT: ISHIMARU, Kaori
; APPLICANT: IGARASHI, Koichi
; APPLICANT: TERADA, Masaaki
; TITLE OF INVENTION: hst-2 MUTEIN, DNA CODING FOR THE SAME AND
; PREPARATION THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/551.171
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/196.001
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: NEUNER, George W
; REGISTRATION NUMBER: 26964
; REFERENCE/DOCKET NUMBER: 44146
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)523-3400
; TELEFAX: (617)523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 171 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-551-171-2

```

Db	44	G I K Q R R L Y C -- N V G I G F H Q L V L P D G R I S T H E E N P Y S L L E I S T V E R G V V S L F G V R S A L F	101
Qy	103	L C Q O P D G A L G S P H F D P E A C S F R E L L E D G Y I N V Y Q S E A H -- G L P L R L P Q ----- K D S P	153
Db	102	V A M N S K G R L Y A T P S F Q -- E E C K F R E T L L P N N Y A N Y E S D L Y O G T I A L S K Y G R V K R G S K V S P	160
Qy	154	N Q D A T S W G P V R F L P	167
Db	161	I M T V T ----- H F L P	169

RESULT 10
US-08-551-171-7
; Sequence 7, Application US/08551171

RESULT 10
 US-08-551-171-7
 ; Sequence 7, Application US/08551171
 ; Patent No. 5679550
 ; GENERAL INFORMATION:
 ; APPLICANT: YOSHIMURA, Koji
 ; APPLICANT: ISHIMARU, Kaori
 ; APPLICANT: IGARASHI, Koichi
 ; APPLICANT: TERADA, Masaaki
 ; TITLE OF INVENTION: hst-2 MUTEIN, DNA CODING FOR THE SAME AND
 ; PREPARATION THEREOF

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

RESULT 10
US-08-551-171-7
; Sequence 7, Application US/08551171
; Patent No. 5679550
; GENERAL INFORMATION:
; APPLICANT: YOSHIMURA, Koji
; APPLICANT: ISHIMARU, Kaori
; APPLICANT: IGARASHI, Koichi
; APPLICANT: TERADA, Masaaki
; TITLE OF INVENTION: hst-2 MUTEIN, D
; TITLE OF INVENTION: PREPARATION THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, J
; ADDRESS: CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0,
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/551,171
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/196,001
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: NEUNER, George W
; REGISTRATION NUMBER: 26964
; REFERENCE/DOCKET NUMBER: 44146
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)523-3400
; TELEFAX: (617)523-6440
; TELE: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-551-171-7

```

	Query Match	15.9%	Score 177.5;	DB 1;	Length 174;
	Best Local Similarity	34.3%;	Pred. No. 2.3e-11;		
	Matches	46;	Mismatches	50;	Indels 17; Gaps
Qy	43	GQVRQRVLYTDDQDTBAHLEIREGTVWGAAHRSPSLELLEKALPKVGVIILGVKASRF	102		
		: : : : : : : : : : :			
Db	47	GIKRRLRYC--NVGIGFHLQLVDGRISGTHEENPYSLLIEISTVERGVVSLEFGVRSALF	104		
Qy	103	LCOQPDGALYGSPIHFDPACSFRELLEDGVNYYQSBAH--GLPLRLPQ-----KDPS	153		
		: : : : : : : : : : : : :			
Db	105	VAMNSKGRLYATPSFO--EECFRFTLLPNKNYNVESDLYOGTYIALSKYGRVKRGSKVSP	163		

QY 154 NQDATSWGVPVRLP 167
Db 164 INTVT-----HFLP 172

RESULT 11

US-08-902-233-7

; Sequence 7, Application US/08902233

; Patent No. 6013784

; GENERAL INFORMATION:

; APPLICANT: YOSHIMURA, Koji

; APPLICANT: ISHIMARU, Kaori

; APPLICANT: IGARASHI, Koichi

; APPLICANT: TERADA, Masaaki

; TITLE OF INVENTION: hst-2 MUTEIN, DNA CODING FOR THE SAME AND

; TITLE OF INVENTION: PREPARATION THEREOF

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &

; ADDRESSEE: CUSHMAN

; STREET: 130 Water Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: US

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/902,233

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/196,001

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: NEUNER, George W

; REGISTRATION NUMBER: 26964

; REFERENCE/DOCKET NUMBER: 44146

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617)523-3400

; TELEFAX: (617)523-6440

; TELEX: 200291 STRE UR

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 174 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-902-233-7

Query Match 15.9%; Score 177.5; DB 3; Length 174;

Best Local Similarity 34.3%; Pred. No. 2.3e-11;

Matches 46; Conservative 21; Mismatches 50; Indels 17; Gaps 5;

QY 43 GQVRQRYLYTDDQDTEAHLEIREDTGVVGAHRSPELLELKALKPGVQILGVKASRF 102

Db 47 GIKRQRLRYC--NVGIGFHLQVLPDGRISGTHEENPYSLEISTVERGVSLFGVRSALF 104

QY 103 LCOQPDGALYGFPHDFPACSFRELLEDGYNVYQSEAH-GLPRLRPQ-----KDSP 153

Db 105 VAMNSKGLRYATPSFQ-ECKFRFETLLPNNYNAYESDLYQGTIYALSKEYGRVKRGSVSP 163

QY 154 NQDATSWGVPVRLP 167

Db 164 INTVT-----HFLP 172

RESULT 12

US-08-551-171-8

; Sequence 8, Application US/08551171
; Patent No. 5679550

; GENERAL INFORMATION:

; APPLICANT: YOSHIMURA, Koji

; APPLICANT: ISHIMARU, Kaori

; APPLICANT: IGARASHI, Koichi

; APPLICANT: TERADA, Masaaki

; TITLE OF INVENTION: hst-2 MUTEIN, DNA CODING FOR THE SAME AND

; TITLE OF INVENTION: PREPARATION THEREOF

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &

; ADDRESSEE: CUSHMAN

; STREET: 130 Water Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: US

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/551,171

; FILING DATE:

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/196,001

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: NEUNER, George W

; REGISTRATION NUMBER: 26964

; REFERENCE/DOCKET NUMBER: 44146

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617)523-3400

; TELEFAX: (617)523-6440

; TELEX: 200291 STRE UR

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 175 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-551-171-8

Query Match 15.9%; Score 177.5; DB 1; Length 175;

Best Local Similarity 34.3%; Pred. No. 2.4e-11;

Matches 46; Conservative 21; Mismatches 50; Indels 17; Gaps 5;

QY 43 GQVRQRYLYTDDQDTEAHLEIREDTGVVGAHRSPELLELKALKPGVQILGVKASRF 102

Db 48 GIKRQRLRYC--NVGIGFHLQVLPDGRISGTHEENPYSLEISTVERGVSLFGVRSALF 105

QY 103 LCOQPDGALYGFPHDFPACSFRELLEDGYNVYQSEAH-GLPRLRPQ-----KDSP 153

Db 106 VAMNSKGLRYATPSFQ-ECKFRFETLLPNNYNAYESDLYQGTIYALSKEYGRVKRGSVSP 164

QY 154 NQDATSWGVPVRLP 167

Db 165 INTVT-----HFLP 173

RESULT 13

US-08-902-233-8

; Sequence 8, Application US/08902233

; Patent No. 6013784

; GENERAL INFORMATION:

; APPLICANT: YOSHIMURA, Koji

; APPLICANT: ISHIMARU, Kaori

; APPLICANT: IGARASHI, Koichi

; APPLICANT: TERADA, Masaaki


```

; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: H095-01A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-441-629-12

```

```

Query Match      15.9%; Score 177.5; DB 1; Length 198;
Best Local Similarity 34.3%; Pred. No. 2.8e-11;
Matches 46; Conservative 21; Mismatches 50; Indels 17; Gaps 5;

QY  43  GQVRQRYLYTDDQDTEAHLEIREDTGVVGAHRSPESELLELKALKPGVIOILGVKASRF 102
    | ||| || : ||: || : || || || || || || || || || || || || || ||
Db  71  GIKRQRRLYC--NVGTGFHLQVLPDGRISGTHEENPYSLLEISTVERGVWSLFGVRSALF 128

QY  103  LCQQPDGALYGSFHPDPEACSPRELLEDGYNVYOSEAH-GLPLRLPO-----KDSP 153
    : ||| :|| | || || || || || || || || || || || || || || || ||
Db  129  VAMNSKGRLYATPSFQ-EECKFRETLLPNNYNAYESDLYQGYIALSKYGRVKRGSKVSP 187

QY  154  NQDATSMGVPVRELP 167
    | |||
Db  188  IMTVT-----HELP 196

```

Search completed: August 6, 2001, 21:05:43
Job time: 1064 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 6, 2001, 20:57:48 ; Search time 64.04 Seconds
(without alignments)
198.798 Million cell updates/sec

Title: US-09-391-861-4
Perfect score: 1115
Sequence: 1 MEMMSRVGTGLGWVRLLA.....SSDPLSMVEPLOGRSPSYAS 210

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0601.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1115	100.0	210	21	AA18636
2	603.5	54.1	153	22	AA18636
3	262.5	23.5	216	20	AA18636
4	262.5	23.5	216	20	AA18636
5	262.5	23.5	216	20	AA18636
6	262.5	23.5	216	21	AA18636
7	262.5	23.5	216	21	AA18636
8	262.5	23.5	216	22	AA18636
9	262.5	23.5	216	22	AA18636
10	262.5	23.5	216	22	AA18636
11	257	23.0	215	20	AA18636

12	256.5	23.0	214	20	AA18636
13	241	21.6	251	22	AA18636
14	207.5	18.6	142	22	AA18636
15	207.5	18.6	213	20	AA18636
16	183	16.4	175	15	AA18636
17	177.5	15.9	146	15	AA18636
18	177.5	15.9	171	15	AA18636
19	177.5	15.9	198	16	AA18636
20	177.5	15.9	198	16	AA18636
21	177.5	15.9	198	16	AA18636
22	177.5	15.9	198	21	AA18636
23	177.5	15.9	198	21	AA18636
24	177.5	15.9	198	21	AA18636
25	177.5	15.9	198	21	AA18636
26	177.5	15.9	207	21	AA18636
27	177.5	15.9	208	13	AA18636
28	177.5	15.9	208	15	AA18636
29	177.5	15.9	208	15	AA18636
30	177.5	15.9	208	22	AA18636
31	177.5	15.9	208	22	AA18636
32	170.5	15.3	239	16	AA18636
33	170.5	15.3	239	16	AA18636
34	170.5	15.3	239	19	AA18636
35	170.5	15.3	239	21	AA18636
36	170.5	15.3	239	21	AA18636
37	170.5	15.3	239	21	AA18636
38	170.5	15.3	239	22	AA18636
39	170.5	15.3	239	22	AA18636
40	170.5	15.3	239	22	AA18636
41	170.5	15.3	239	22	AA18636
42	169.5	15.2	240	21	AA18636
43	165.5	14.8	140	15	AA18636
44	165.5	14.8	148	13	AA18636
45	165.5	14.8	148	15	AA18636

ALIGNMENTS

RESULT 1

AA18636

ID AA18636 standard; Protein; 210 AA.

AC AA18636;

DT 22-JAN-2001 (first entry)

XX Amino acid sequence of murine fibroblast growth factor (FGF)-21.

XX PD10-VEGFuc; gene delivery vector; eye disease; neovascular disease;
XX neurotrophic factor; anti-angiogenic factor; eye disease; glaucoma;
XX macular degeneration; diabetic retinopathy; retinitis pigmentosa;
XX inherited retinal degeneration; surgery-induced retinopathy;
XX retinal detachment; photic retinopathy; toxic retinopathy;
XX trauma-induced retinopathy; wet age related macular degeneration;
XX ARMD; retinopathy; fibroblast growth factor-20; FGF-20.

OS Mus sp.

XX WO200054813-A2.

XX 21-SEP-2000.

XX 15-MAR-2000; 2000WO-US07062.

XX 15-MAR-1999; 99US-0124460.

XX 06-JAN-2000; 2000US-0174984.

XX (CHIR) CHIRON CORP.

XX (REGC) UNIV CALIFORNIA.

XX Manning WC, Dwarki VJ, Rendahl K, Zhou S, McGee LH, Lau D;

XX Flannery JG, Miller S, Wang F, Di Polo A;

```
XX WPI: 2000-618862/59.
DR N-PSDB; AAA75631.
XX
XX Treating or preventing eye diseases or inhibiting neovascular disease
PT of the eye, comprises intraocularly administering a gene delivery
PT vector that directs expression of neurotrophic factors or
PT anti-angiogenic factors -
XX
XX Disclousure; Fig 30; 86pp; English.
XX
XX The present sequence represents a murine fibroblast growth factor (FGF)
CC -21. FGF-20 is expressed using a gene delivery vector of the invention.
CC Vectors of the invention are used for treating or preventing eye
CC diseases, or inhibiting neovascular disease of the eye. The gene
CC delivery vector directs the expression of one or more neurotrophic
CC factors, or anti-angiogenic factors, such that the disease of the eye
CC is prevented or treated. The gene delivery vectors are useful for
CC treating or preventing diseases of the eye such as macular degeneration,
CC diabetic retinopathy, inherited retinal degeneration such as retinitis
CC pigmentosa, glaucoma, surgery-induced retinopathy, retinal detachment,
CC photic retinopathy, toxic retinopathy or trauma-induced retinopathy and
CC for inhibiting neovascular diseases such as wet age related macular
CC degeneration (ARM) or retinopathy of prematurity.
XX
XX Sequence 210 AA;
SQ
Query Match 100.0%; Score 1115; DB 21; Length 210;
Best Local Similarity 100.0%; Pred. No. 4.8e-102;
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEWMSRVCTGLWVRLLLAVFLGVOAYPIPDSSPLLOFGGVQRQRYLYTDDDDPTEA 60
DB 1 mewmrsrvctglwvrlilavllgvyqaypdpssplllqf99qvrqrylytdddqdea 60
QY 61 HLEIRDTGVVGAHRSPELLEKALPGVVIQILGVKASRFLCQOPDGCALYSGSPHFDPPE 120
DB 61 hleiredgtvgaahrspelellkalkpgvviqlgvkasrfllcqpqdgalygsphfdpe 120
QY 121 ACSFRELLEDGYNVYQSEAHGLPLRLPKOKSPNQDQATSWGPVRLPMRGLLHEPDQAG 180
DB 121 acsfrellledgynvyqseahglplrlpkdpsnqdatSWGpvrlpmpgllhepqdqag 180
QY 181 FLPPPEPDVGSDDLPMVEPLQGRSPSYAS 210
DB 181 flppeppdvgsddplsmveplqgrspsyas 210
RESULT 2
ID AAB73069 standard; Protein; 153 AA.
AC AAB73069;
XX
XX 17-MAY-2001 (first entry)
DT
DE Human fibroblast growth factor 19 homologue FGF19X.
XX
XX Human; fibroblast growth factor 19; FGF19X; embryonic development;
KW proliferative disorder; cancer; restenosis; psoriasis;
KW rheumatoid arthritis; Dupuytren's contracture.
XX
XX Homo sapiens.
OS
XX WO200118209-A1.
PN
XX 15-MAR-2001.
PD
XX 08-SEP-2000; 2000WO-US24863.
PF
XX 10-SEP-1999; 99US-0153303.
PR
XX 03-APR-2000; 2000US-0194246.
PR
```

```
PR 19-JUL-2000; 2000US-0619251.
XX (CURA-) CURAGEN CORP.
XX
XX Shimkets RA, Vernet C, Burgess C, Fernandes E, Taupier R;
PI Quinn KE, Spyteck KA, Rastelli L, Herrmann JL;
XX
XX WPI: 2001-218559/22.
DR N-PSDB; AAF76715.
XX
XX Fibroblast growth factor-19X polypeptides and polynucleotides useful
PT for diagnosis, prevention, treatment of proliferative, differentiative,
PT tumorigenic disorders such as tumor, psoriasis, rheumatoid arthritis -
XX
XX Claim 11; Page 7; 97pp; English.
XX
XX The present invention describes the protein and coding sequences of the
CC human fibroblast growth factor (FGF) 19 homologue FGF19X, the former of
CC which is shown here. The sequences can be used in the prognosis and
CC treatment of proliferative diseases such as cancer, restenosis,
CC psoriasis, rheumatoid arthritis and Dupuytren's contracture, as well as
CC to stimulate cell growth for treating neurological disorders such as
CC Alzheimer's disease.
XX
XX Sequence 153 AA;
SQ
Query Match 54.1%; Score 603.5; DB 22; Length 153;
Best Local Similarity 86.9%; Pred. No. 7.7e-52;
Matches 119; Conservative 5; Mismatches 12; Indels 1; Gaps 1;
QY 12 GLWVRLLLAVFLGVOAYPIPDSSPLLOFGGVQRQRYLYTDDDDQTEAHLEIREDTGVV 71
DB 12 glwvs-vlaglllgacqahpdpssplllqf99qvrqrylytddaqteahleiredgtvg 70
QY 72 GAAHRSPELLEKALPGVVIQILGVKASRFLCQOPDGCALYSGSPHFDPPEACSFRELLED 131
DB 71 gaadqspesllqalkpgvviqlgvksrfllcqpqdgalygslhfdpeacsfrellled 130
QY 132 GYNVYQSEAHGLPLRLP 148
DB 131 gynyvqseahglplhlp 147
RESULT 3
ID AAY08583
XX AAY08583 standard; Protein; 216 AA.
XX
XX AC AAY08583;
XX
XX 05-AUG-1999 (first entry)
DT
DE Human PRO533 protein fragment.
XX
XX PRO533; FGF-19; fibroblast growth factor; human; diagnosis; treatment;
KW tumour; neoplastic cell growth; cell proliferation; tumorigenesis; cancer;
KW autocrine signalling.
XX
XX Homo sapiens.
OS
XX WO9927100-A1.
PN
XX 03-JUN-1999.
PD
XX 25-NOV-1998; 98WO-US25190.
PF
XX 21-SEP-1998; 98US-0158432.
PR
XX 25-NOV-1997; 97US-0066840.
XX
XX (GETH ) GENENTECH INC.
XX
XX Botstein D, Goddard A, Gurney AL, Hillan KJ, Lawrence DA;
PI Roy MA;
PR
```


XX		
PR	25-NOV-1997;	97US-0066840.
PR	17-SEP-1997;	97US-0059113.
PR	17-SEP-1997;	97US-0059115.
PR	17-SEP-1997;	97US-0059117.
PR	17-SEP-1997;	97US-0059119.
PR	17-SEP-1997;	97US-0059121.
PR	17-SEP-1997;	97US-0059122.
PR	17-SEP-1997;	97US-0059184.
PR	18-SEP-1997;	97US-0059263.
PR	15-OCT-1997;	97US-0059266.
PR	17-OCT-1997;	97US-0062125.
PR	17-OCT-1997;	97US-0062285.
PR	21-OCT-1997;	97US-0063486.
PR	24-OCT-1997;	97US-0062814.
PR	24-OCT-1997;	97US-0062816.
PR	24-OCT-1997;	97US-0063045.
PR	24-OCT-1997;	97US-0063120.
PR	24-OCT-1997;	97US-0063121.
PR	24-OCT-1997;	97US-0063127.
PR	27-OCT-1997;	97US-0063128.
PR	27-OCT-1997;	97US-0063329.
PR	28-OCT-1997;	97US-0063327.
PR	28-OCT-1997;	97US-0063341.
PR	28-OCT-1997;	97US-0063342.
PR	28-OCT-1997;	97US-0063344.
PR	28-OCT-1997;	97US-0063349.
PR	28-OCT-1997;	97US-0063550.
PR	29-OCT-1997;	97US-0063564.
PR	29-OCT-1997;	97US-0063435.
PR	29-OCT-1997;	97US-0063704.
PR	29-OCT-1997;	97US-0063732.
PR	29-OCT-1997;	97US-0063738.
PR	29-OCT-1997;	97US-0063734.
PR	29-OCT-1997;	97US-0064215.
PR	29-OCT-1997;	97US-0063735.
PR	31-OCT-1997;	97US-0063870.
PR	31-OCT-1997;	97US-0064103.
PR	03-NOV-1997;	97US-0064248.
PR	07-NOV-1997;	97US-0064809.
PR	12-NOV-1997;	97US-0065186.
PR	17-NOV-1997;	97US-0065846.
PR	18-NOV-1997;	97US-0065693.
PR	21-NOV-1997;	97US-0066120.
PR	21-NOV-1997;	97US-0066364.
PR	24-NOV-1997;	97US-0066772.
PR	24-NOV-1997;	97US-0066466.
PR	24-NOV-1997;	97US-0066770.
PR	24-NOV-1997;	97US-0066511.
PR	24-NOV-1997;	97US-0066453.
XX		
PA	(GETH) GENENTECH INC.	
XX		
PI	Chen J, Goddard A, Gurney AL, Pennica D, Wood WI, Yuan J;	
XX		
DR	WPI; 1999-229533/19.	
DR	N-PSDB; AAX52224.	
XX		
XX	New isolated human genes and polypeptides used in, e.g. treatment of	
PT	gastrointestinal ulceration	
XX		
PS	Claim 12; Fig 22; 320pp; English.	
XX		
CC	AA13344-403 represent secreted and transmembrane human proteins.	
CC	The cDNA sequences are obtained from cDNA libraries, prepared from	
CC	fetal lung, fetal kidney, fetal brain, fetal liver and fetal retina.	
CC	The encoded polypeptides have specific uses based on their homology to	
CC	known polypeptides, e.g. PRO211 and PRO217 can be used for disorders	
CC	associated with the preservation and maintenance of gastrointestinal	
CC	mucosa and the repair of acute and chronic mucosal lesions	
CC	(e.g. enterocolitis, Zollinger-Ellison syndrome, gastrointestinal	

CC		with abnormal keratinocyte differentiation (e.g. psoriasis, epithelial cancers such as lung squamous cell carcinoma of the vulva and gliomas),
CC		potent effects on cell growth and development, diseases related to
CC		growth or survival of nerve cells including Parkinson's disease,
CC		Alzheimer's disease, ALS, neuropathies or cancer. PRO265 can be used
CC		for fibromodulin, e.g. for reducing dermal scarring. PRO264 can be used
CC		as a target for anti-tumor drugs. PRO333 may be used in the treatment
CC		of Usher Syndrome or Atrophia areata; PRO269 can be used as an
CC		anti-thrombotic agent; PRO287 polypeptides and portions may have
CC		therapeutic applications in wound healing and tissue repair; PRO317 can
CC		be used for treating problems of the kidney, uterus, endometrium, blood
CC		vessels, or related tissue, e.g. in the heart of genital tract.
XX		
SS		
QS		
	Sequence	216 AA;
	Query Match	23.5%; Score 262.5; DB 20; Length 216;
	Best Local Similarity	34.7%; Pred. No. 4.8e-18;
	Matches	75; Conservative 33; Mismatches 95; Indels 13; Gaps 8;
QY	4	MRSRVGTGLVRRLLAVFLLVGVQAYPIPDSPSLQF--GGQVRORYLYTDDQD-RFA 60 : : : : : : : : : : : : : : : : : : : : : :
Db	1	mrgscvvhvwf--laglwlavagrpiafadagphvygwdplrlrhlytsgphglssc 58
QY	61	HLEIRDEGTVVGAAHRSPESLLELKALKPGVIQTILGVKASRFLCQQPDGALYGSPHFDPE 120
Db	59	fllradgvvdccargqsahsleikavalrtvaikgvhsrvylcmgadqkmagllqysee 118
QY	121	ACSFRELLEDGYNYVQSEAHGLPLRLPQ-KDSPNQDATSWGVP-RFLPM-PCLLHEPOD 177 : :
Db	119	dcafeeeirpdgynyvrsekhlrplvsissakgrglyknrgflplshflpmlpmvpeeped 178
QY	178	QAQFLTPPE----PPDVGSSDPLSMVEPLQG-RSPSY 208
Db	179	lrghlesdmffspletdmsdpfgltvgleavrspsf 214
RESULT	5	
AAY05280	ID	AAY05280 standard; Protein; 216 AA.
XX	XX	AAY05280;
AC	XX	
DT	DT	22-JUN-1999 (first entry)
DE	DE	FGF-8 homologue PRO533.
XX	XX	
KW	KW	Antibody; PRO187; PRO533; PRO214; PRO240; PRO230; PRO261; PRO246;
KW	KW	EBAF-2; inhibitor; tumour growth; cancer; EGF-like homologue;
XX	XX	FGF-8 homologue.
OS	OS	Homo sapiens.
XX	XX	
PX	PX	WO9914327-A2.
XX	XX	
PD	PD	25-MAR-1999.
XX	XX	
XX	XX	10-SEP-1998; 98WO-US1882A.
XX	XX	
PR	PR	25-NOV-1997; 97US-0066840.
PR	PR	17-SEP-1997; 97US-0059114.
PR	PR	17-SEP-1997; 97US-0059117.
PR	PR	18-SEP-1997; 97US-0059263.
PR	PR	15-OCT-1997; 97US-0062125.
PR	PR	17-OCT-1997; 97US-0062285.
PR	PR	17-OCT-1997; 97US-0062287.
PR	PR	24-OCT-1997; 97US-0062816.
PR	PR	25-OCT-1997; 97US-0063704.
XX	XX	
PA	PA	(GETH) GENENTECH INC.
XX	XX	
PI	PI	Beth Stein D, Goddard A, Gurney A, Hillan K, Lawrence DA,
PI	PI	'Rev Mr., Wood WI;

CC acid encoding the antibodies are used in the production of the
 CC antibodies. The antibodies and nucleic acids encoding them are used for
 CC diagnosing a tumour in a mammal. The antibodies are used for inhibiting
 CC the growth of tumour cells and identifying compounds that inhibit a
 CC biological or immunological activity of and/or expression of a PRO187,
 CC PRO533, PRO214, PRO240, PRO211, PRO230, PRO261, PRO246 or PRO317
 CC polypeptide. The antibody can be used in antibody dependent enzyme
 CC mediated prodrug therapy (ADEPT) by conjugating the antibody to a
 CC prodrug-activating enzyme which converts a prodrug to an anti-cancer
 CC drug. The antibodies can be fluorescently labelled and monitored by light
 CC microscopy, flow cytometry or fluorimetry for diagnosis and prognosis of
 CC tumours.

XX Sequence 216 AA;

Query Match 23.5%; Score 262.5; DB 21; Length 216;
 Best Local Similarity 34.7%; Pred. No. 4.8e-18;
 Matches 75; Conservative 33; Mismatches 95; Indels 13; Gaps 8;

QY 4 MRSRVGTGLGLWRLLLAVFLGVYQAYPIPDSSPLLQF--GGQVRORYLYTDDQD-TEA 60
 DB 1 mrsrgvvvhwii--laglwlavagrplafsdagphvnygwdpirlrhlytsghglssc 58
 QY 61 HLEIREDTGVVGAHRSPELLEKALKPGVQIQLGVKASRFLCQPDGALYGSHPFDPE 120
 DB 59 flriradgvvdvcargqsahsleikavalrtvaikgvsrvylcmgadmkgmgllysee 118
 QY 121 ACSFRELLLEDGNYVQSEAHGLPLRLPQ-KDSPNQDATSWGPV-RFLPW-PGLLHEPDQ 177
 DB 119 dcafeeeirpdgynyvyrsekhrlpvsalsakqrlgknygfplshflpmlpmvpeeped 178
 QY 178 QAGFLPPE---PPDVGSSDPLSMVEPLQG-RSPSY 208
 DB 179 lrglhlesdmfsspletddmpfgivtgleavrspsf 214

RESULT 8

AAB68593
 ID AAB68593 standard; Protein; 216 AA.

XX AAB68593;

DT 27-APR-2001 (first entry)

XX PRO533.

XX Cytostatic; PRO protein; tumour; cancer.

XX Homo sapiens.

XX WO200105836-A1.

XX 25-JAN-2001.

XX 20-DEC-1999; 99WO-US30999.

XX 20-JUL-1999; 99US-0144758.

XX 26-JUL-1999; 99US-0145698.

XX 08-SEP-1999; 99WO-US20594.

XX 13-SEP-1999; 99WO-US20944.

XX 15-SEP-1999; 99WO-US21090.

XX 05-OCT-1999; 99WO-US23089.

XX 29-NOV-1999; 99WO-US28214.

XX 30-NOV-1999; 99WO-US28313.

XX 02-DEC-1999; 99WO-US28564.

XX (GETH) GENENTECH INC.

XX Botstein D, Goddard A, Gurney AL, Hillan KJ, Roy MA, Wood WI;

XX WPI; 2001-091968/10.

DR N-PSDB; AAF58498.

XX New antibody that binds to a PRO polypeptide, e.g. PRO187 and PRO533,
 PT useful for diagnosing and treating cancers -
 XX Claim 61; Fig 4; 196pp; English.

XX The present invention relates to PRO proteins and coding sequences. The
 CC present sequence is one such PRO protein. It was found that the PRO genes
 CC are amplified in the genome of tumour cells. The gene amplification is
 CC expected to be associated with the overexpression of the gene product and
 CC contributes to tumourigenesis. Therefore, antagonists of PRO proteins are
 CC useful for the treatment of benign or malignant tumours, leukaemias,
 CC lymphoid malignancies and other disorders such as neuronal, glial,
 CC astrocytal, hypothalamic, glandular, epithelial, inflammatory and
 CC immunologic disorders.

XX Sequence 216 AA;

Query Match 23.5%; Score 262.5; DB 22; Length 216;
 Best Local Similarity 34.7%; Pred. No. 4.8e-18;
 Matches 75; Conservative 33; Mismatches 95; Indels 13; Gaps 8;

QY 4 MRSRVGTGLGLWRLLLAVFLGVYQAYPIPDSSPLLQF--GGQVRORYLYTDDQD-TEA 60
 DB 1 mrsrgvvvhwii--laglwlavagrplafsdagphvnygwdpirlrhlytsghglssc 58
 QY 61 HLEIREDTGVVGAHRSPELLEKALKPGVQIQLGVKASRFLCQPDGALYGSHPFDPE 120
 DB 59 flriradgvvdvcargqsahsleikavalrtvaikgvsrvylcmgadmkgmgllysee 118
 QY 121 ACSFRELLLEDGNYVQSEAHGLPLRLPQ-KDSPNQDATSWGPV-RFLPW-PGLLHEPDQ 177
 DB 119 dcafeeeirpdgynyvyrsekhrlpvsalsakqrlgknygfplshflpmlpmvpeeped 178
 QY 178 QAGFLPPE---PPDVGSSDPLSMVEPLQG-RSPSY 208
 DB 179 lrglhlesdmfsspletddmpfgivtgleavrspsf 214

RESULT 9

AAB31201

ID AAB31201 standard; Protein; 216 AA.

XX AAB31201;

XX 20-APR-2001 (first entry)

XX Amino acid sequence of human polypeptide PRO553.

XX Human; secreted protein; transmembrane protein; PRO196; PRO444; PRO183;
 KW PRO185; PRO210; PRO215; PRO217; PRO242; PRO288; PRO365; PRO1361; PRO1308;
 KW PRO1183; PRO1272; PRO1419; PRO4999; PRO7170; PRO248; PRO353; PRO1318;
 KW PRO1600; PRO9940; PRO533; PRO301; PRO187; PRO337; PRO1411; PRO4356;
 KW PRO246; PRO265; PRO941; PRO10096; PRO6003; PRO6004; PRO350; PRO2630;
 KW PRO6309; cell death; genetic disorder; transgenic animal; gene therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Peptide 1..22 /note= "signal peptide"

XX Modified-site 15..21 /note= "N-myristoylation site"

XX Binding-site 48..59 /note= "prokaryotic membrane lipoprotein lipid attachment site"

XX Modified-site 54..60 /note= "N-myristoylation site"

XX Modified-site 66..72 /note= "N-myristoylation site"

XX Modified-site 78..82 /note= "casein kinase II phosphorylation site"

XX PA (GETH) GENENTECH INC.

XX Botstein D, Goddard A, Gurney AL, Hillan KJ, Lawrence DA;

PI Roy MA;

XX WPI; 1999-347718/29.

XX Nucleic acid encoding fibroblast growth factor - 19, useful for the

PT diagnosis, prevention and treatment of cancers

PT Disclosure; Fig 3; 88pp; English.

PS This invention describes a novel human fibroblast growth factor, PRO533,

XX also known as fibroblast growth factor-19 (FGF-19). The nucleic acids,

CC methods and PRO533 polypeptides disclosed may be used in the diagnosis

CC and treatment of tumours and/or conditions characterized by modulation of

CC PRO533 expression, or in the preparation of compositions for such

CC therapies. These compositions and methods may be used in the diagnosis

CC and treatment of neoplastic cell growth and proliferation in mammals

CC (especially humans). The invention is based on the identification of

CC genes that are amplified in the genome of tumour cells. Such gene

CC amplification is expected to be associated with the over expression of

CC the gene product and contribute to tumourgenesis and/or autocrine

CC signalling. Accordingly, the proteins encoded by the amplified genes are

CC believed to be useful targets for the diagnosis and/or treatment of

CC certain cancers and may act as predictors of the prognosis for tumour

XX treatments.

XX Sequence 214 AA;

XX Query Match 23.0%; Score 256.5; DB 20; Length 214;

XX Best Local Similarity 34.6%; Pred. No. 1.9e-17;

XX Matches 73; Conservative 33; Mismatches 94; Indels 11; Gaps 7;

QY 9 GTGLGWRLLAVFLGCVQAYPIPPSSPLLQF--GGVQRQRYLYTDDQD--TEAHLER 65

Db 2 gcvvnhwiliaglwladgrlafsdagphvhygdprlrlhlytsgphglsscfllr 61

QY 66 EDGTVVGAHRSPESLLELKALPKPGVITQLGVKASRFLCQPDGALYSPHFDPEACSPR 125

Db 62 adgvdcargqsaahsleikavalrtvaikghsvrylcmgagdkmqgllqyseedcafe 121

QY 126 ELLEDEGVNYQSEAHGLPLRLPQ-KDSPNQDATSNQPV-RFLPM-PGLLHPEPDQAGFL 182

Db 122 eairpdgynvyrrsekhrplvslssakgrqlyknrgflplshfplmpmypepedlrgl 181

QY 183 PPE---PPDVGSSDPLSNVEPLOG-RSPSY 208

Db 182 esdmfsspletmdmpfiglvtgieavrspsf 212

RESULT 13

AAB65297

ID AAB65297 standard; Protein; 251 AA.

XX AAB65297;

XX 02-APR-2001 (first entry)

XX Human PRO9828 protein sequence SEQ ID NO:511.

XX Human; secreted and transmembrane protein; PRO; cytostatic;

KW cell death; cancer; chromosomal mapping; gene mapping; tissue typing;

KW diagnostic assay.

XX Homo sapiens.

OS WO200073454-A1.

XX 07-DEC-2000.

XX

PF 30-MAR-2000; 2000WO-US08439.

XX 02-JUN-1999; 99WO-US12252.

PR 23-JUN-1999; 99US-0141037.

PR 07-JUL-1999; 99US-0143048.

PR 20-JUL-1999; 99US-0144738.

PR 26-JUL-1999; 99US-0145698.

PR 28-JUL-1999; 99US-0146222.

PR 17-AUG-1999; 99US-0149396.

PR 15-SEP-1999; 99WO-US21090.

PR 15-SEP-1999; 99WO-US21547.

PR 08-OCT-1999; 99US-0158663.

PR 30-NOV-1999; 99WO-US28313.

PR 01-DEC-1999; 99WO-US28301.

PR 16-DEC-1999; 99WO-US30095.

PR 20-DEC-1999; 99WO-US30911.

PR 05-JAN-2000; 2000WO-US00219.

PR 06-JAN-2000; 2000WO-US00376.

PR 11-FEB-2000; 2000WO-US03565.

PR 18-FEB-2000; 2000WO-US04341.

PR 22-FEB-2000; 2000WO-US04414.

PR 24-FEB-2000; 2000WO-US04914.

PR 24-FEB-2000; 2000WO-US05004.

PR 02-MAR-2000; 2000WO-US05841.

PR 15-MAR-2000; 2000WO-US06884.

PR 20-MAR-2000; 2000WO-US07377.

XX (GETH) GENENTECH INC.

XX Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;

PI Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;

PI Grimaldi CJ, Gurney AL, Kljavin IJ, Napier MA, Pan J, Paoni NF;

PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;

PI Zhang Z;

XX WPI; 2001-032160/04.

DR N-PSDB; AAF44266.

XX PRO polynucleotides used to produce polypeptides used to target

PT bioactive molecules such as toxins, radiolabels or antibodies, to

XX specific cells, to cause targeted cell death -

PS Claim 12; Fig 324; 935pp; English.

XX The present invention describes human secreted and transmembrane PRO

CC proteins. The PRO proteins have cytostatic activity. The PRO proteins

CC can be used for targeted delivery of bioactive molecules, such as

CC toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide

CC sequences, and their fragments, can be used as hybridisation probes, in

CC chromosomal and gene mapping, and in the generation of anti-sense RNA

CC and DNA. They may also be used to produce transgenic animals which are

CC used to develop and screen therapeutically useful reagents. The PRO

CC nucleotide and protein sequence can be used for tissue typing and in

CC treating cancer. Anti-PRO antibodies can be used in diagnostic assays.

CC AAF44270 to AAF44470 represent PCR primers and hybridisation probes used

CC in the isolation of human PRO sequences. AAF44087 to AAF44269 and

CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein

CC sequences given in the exemplification of the present invention.

XX Sequence 251 AA;

SQ

Query Match 21.6%; Score 241; DB 22; Length 251;

Best Local Similarity 32.3%; Pred. No. 7.7e-16;

Matches 71; Conservative 38; Mismatches 71; Indels 40; Gaps 8;

QY 11 LGLWVRLLLAVFLGCVQAYPIPPSSPLL--QFGGVQRQRYLYTDDQDTEAHLEREDG 68

Db 6 lrlwvcalscvcsmsvray--pnaspllgsswglli---hlytatarns-yhliqhkng 59

QY 69 TVVGAHRSPESLLELKALPKPGVITQLGVKASRFLCQPDGALYSPHFDPEACSFRELL 128

Db 60 hvdgaphqtiysalmirsedagfvitgmrvrylcmfgrnifgshyfdpencrfghqt 119

Db 129 dclgynyrsmkhlhlfiqa-kpreqlqdqkpsnfip---vfhrsffetgdqlrskmf 184
Qy 182 LPPEPPDVGSDDLPLSMVEPLQG--RSPSY 208
Db 185 slplesd--smdpfrmvedvhlvkpsf 211

Search completed: August 6, 2001, 20:57:49
Job time: 9721 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 6, 2001, 22:37:53 ; Search time 2566.06 Seconds
(without alignments)
3912.056 Million cell updates/sec

Title: US-09-391-861-3
Perfect score: 649
Sequence: 1 atggaatgatgagatctag.....ctgactcttcctgaatcta 649

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 segs, 773874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_cm:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_p11:*
13: gb_p12:*
14: gb_p13:*
15: gb_p14:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_hugo_hum:*
20: em_hugo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_cm:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_p1:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_v1:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_v11:*
59: gb_v12:*
60: gb_vtq1:*
61: gb_vtq2:*
62: gb_vtq3:*
63: gb_vtq4:*
64: gb_vtq5:*
65: gb_vtq6:*
66: gb_vtq7:*
67: gb_vtq8:*
68: gb_vtq9:*
69: gb_vtq10:*
70: gb_vtq11:*
71: gb_vtq12:*
72: gb_vtq13:*
73: gb_vtq14:*
74: gb_vtq15:*
75: gb_vtq16:*
76: gb_vtq17:*
77: gb_vtq18:*
78: gb_vtq19:*
79: gb_vtq20:*
80: gb_vtq21:*
81: gb_vtq22:*
82: gb_vtq23:*
83: gb_vtq24:*
84: gb_vtq25:*
85: gb_vtq26:*
86: gb_vtq27:*
87: gb_vtq28:*
88: gb_vtq29:*
89: gb_vtq30:*
90: gb_vtq31:*
91: gb_vtq32:*
92: gb_vtq33:*
93: gb_vtq34:*
94: gb_vtq35:*
95: gb_vtq36:*
96: gb_vtq37:*
97: gb_vtq38:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	633	97.5	633	94	AB025718 Mus muscu
2	400.2	61.7	630	85	AB021975 Homo sapi
3	297.4	45.8	215734	75	AC073828 Mus muscu
4	296.6	45.7	9257	94	AF214655 Mus muscu
5	283.8	43.7	514	10	AX097639 Sequence
6	197.2	30.4	170538	68	AC024740 Homo sapi
7	195.6	30.1	110000	60	AC008749 Homo sapi
8	195.6	30.1	110000	60	AC008749_1 Continuation (2 of

C	9	194	29.9	4033	85	AB006136	Homo sapi
C	10	194	29.9	37402	87	AC009002	Homo sapi
C	11	135.2	20.8	227949	70	AC026803	Homo sapi
C	12	68.4	10.5	756	85	AB037973	Homo sapi
C	13	68.4	10.5	1612	89	AF263537	Homo sapi
C	14	57.8	8.9	756	94	AB037889	Homo sapi
C	15	57.8	8.9	1559	94	AF263536	Mus muscu
C	16	54.2	8.4	651	85	AB018122	Mus muscu
C	17	54.2	8.4	2137	10	AX076894	Homo sapi
C	18	54.2	8.4	2137	10	AX076894	Homo sapi
C	19	54.2	8.4	2137	10	AX076894	Homo sapi
C	20	47.8	7.4	2291	8	AF110400	Homo sapi
C	21	43.8	6.7	1824	94	AF007268	Mus muscu
C	22	42.4	6.5	676	12	AF115191	Phaseolus
C	23	42.4	6.5	676	12	AF115191	Phaseolus
C	24	41.4	6.4	677	12	AF115200	Phaseolus
C	25	40.6	6.3	161903	67	AC021996	Homo sapi
C	26	40.6	6.3	166620	68	AC024612	Homo sapi
C	27	40.6	6.3	173765	88	AC022007	Homo sapi
C	28	40.6	6.3	173793	60	AC007783	Homo sapi
C	29	40.6	6.3	191834	69	AC026196	Homo sapi
C	30	40.6	6.3	197360	65	AC018829	Homo sapi
C	31	40.6	6.3	216905	87	AC018809	Homo sapi
C	32	39.4	6.1	11048	92	HS889J22B	Human DNA
C	33	39.4	6.1	155622	84	HS889J22	Human DNA
C	34	38.4	5.9	139290	73	AC068096	Homo sapi
C	35	37.6	5.8	796	76	AC080379	Giardia
C	36	37.4	5.8	190394	76	AC079886	Rattus no
C	37	37.2	5.7	116441	79	AL162257	Homo sapi
C	38	37.2	5.7	126141	80	AL356300	Homo sapi
C	39	37.2	5.7	184467	74	AC073241	Homo sapi
C	40	37	5.7	2426	10	I05274	Sequence 2
C	41	37	5.7	2426	10	I08636	Sequence 1
C	42	37	5.7	3398	93	HSEKPG	Human gene
C	43	37	5.7	3401	10	I08406	Sequence 1
C	44	37	5.7	110195	79	AL162741	Homo sapi
C	45	36.8	5.7	196003	76	AC079419	Mus muscu

ALIGNMENTS

RESULT 1
LOCUS AB025718 633 bp mRNA ROD 11-JUL-2000
DEFINITION Mus musculus mRNA for FGF-21, complete cds.
ACCESSION AB025718
VERSION AB025718.1 GI:9049446
KEYWORDS FGF-21.
SOURCE Mus musculus cDNA to mRNA.
ORGANISM Mus musculus

REFERENCE 1 (sites)
AUTHORS Nishimura,T., Nakatake,Y., Konishi,M. and Itoh,N.
TITLE Identification of a novel FGF, FGF-21, preferentially expressed in the liver(1)
JOURNLS Blochim. Biophys. Acta 1492 (1), 203-206 (2000)
MEDLINE 20461777
2 (bases 1 to 633)
AUTHORS Itoh,N.
REFERENCE Direct Submission
JOURNLS Submitted (02-APR-1999) to the DDBJ/EMBL/Genbank databases
Nobuyuki Itoh, Kyoto University, Graduate School of Pharmaceutical Sciences, Department of Genetic Biochemistry, Yoshida-Shimoadachi, Sakyo, Kyoto, Kyoto 606-8501, Japan
(E-mail:itohnobu@pharm.kyoto-u.ac.jp, Tel:81-75-753-4540, Fax:81-75-753-4600)

FEATURES
SOURCE
1. .633
Location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
1. .633
gene

CDS
/gene="FGF-21"
1. 633
/gene="FGF-21"
/codon_start=1
/product="FGF-21"
/protein_id="BA09416.1"
/db_xref="GI:9049447"
/translation="MEMMSRYGTGLWIRLLAVELGVQAVIPDPSPLQFGGQ
VQRIYXMDPDTEAHLEIRPDGVNVAARSPSLLEIRAKRGVQIIGVARSR
LCOQPDGALYGSFHEPDEACSFRELLLEDGIVYQSEHGLPLRPQDSFNQNTSM
GVRFLLPMFGLLHEPDQAGFLPPPPVQSSDPLSMVEPLQGRSTAS"
BASE COUNT 136 a 198 c 173 g 126 t
ORIGIN

Query Match 97.5%; Score 633; DB 94; Length 633;
Best Local Similarity 100.0%; Pred. No. 6,2e-147;
Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	atggaatgataagataagatgagacccctgagactgtagctgctgctgct	60
DB	1	atggaatgataagataagatgagacccctgagactgtagctgctgctgct	60
QY	61	gtctctctgctgaggggtctcaacaagatacccatctgactcagccctctcag	120
DB	61	gtctctctgctgaggggtctcaacaagatacccatctgactcagccctctcag	120
QY	121	tttgagggtcaagtcgcgcagagagtgatctctcaacagatgagcaacagatcagaagc	180
DB	121	tttgagggtcaagtcgcgcagagagtgatctctcaacagatgagcaacagatcagaagc	180
QY	181	caactgagagatcagagagagatgagatgagagagagagagagagagagagagagagag	240
DB	181	caactgagagatcagagagagatgagatgagagagagagagagagagagagagagagag	240
QY	241	ctctctgagatcagaacgcttgaaagcagagagagagagagagagagagagagagagag	300
DB	241	ctctctgagatcagaacgcttgaaagcagagagagagagagagagagagagagagagag	300
QY	301	agggtctcttggcaacaagcagagatgagatctctcaagatgagatgagatgagatgagatg	360
DB	301	agggtctcttggcaacaagcagagatgagatctctcaagatgagatgagatgagatgagatg	360
QY	361	gacctgagatcagagagagatgagatgagatgagatgagatgagatgagatgagatgagatg	420
DB	361	gacctgagatcagagagagatgagatgagatgagatgagatgagatgagatgagatgagatg	420
QY	421	catggcctgcccctgctgctgctcctcagagagagatccccaacagagatgcaacatctgg	480
DB	421	catggcctgcccctgctgctgctcctcagagagagatccccaacagagatgcaacatctgg	480
QY	481	gagacgtgagctctctctgcccagatgagagagatgagagagagagagagagagagagagag	540
DB	481	gagacgtgagctctctctgcccagatgagagagatgagagagagagagagagagagagagag	540
QY	541	ttctctgcccag	600
DB	541	ttctctgcccag	600
QY	601	ttacaggagcgaagccagctatgctgctgca 633	
DB	601	ttacaggagcgaagccagctatgctgctgca 633	

RESULT 2
LOCUS AB021975 630 bp mRNA PRI 03-AUG-2000
DEFINITION Homo sapiens mRNA for FGF-21, complete cds.
ACCESSION AB021975
VERSION AB021975.1 GI:9049444
KEYWORDS new FGF.
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (sites) Nishimura,T., Nakatake,Y., Konishi,M. and Itoh,N. Identification of a novel FGF, FGF-21, preferentially expressed in the liver(1) Biochim. Biophys. Acta 1492 (1), 203-206 (2000) 20461777 Itoh,N. 2 (bases 1 to 630) Direct Submission Submitted (29-DEC-1998) to the DDBJ/EMBL/GenBank databases. Nobuyuki Itoh, Kyoto University, Graduate School of Pharmaceutical Sciences, Department of Genetic Biochemistry, Yoshida-Shimoadachi, Sakyo, Kyoto 606-8501, Japan (E-mail: itohnobu@pharm.kyoto-u.ac.jp, Tel: 81-75-753-4540, Fax: 81-75-753-4600) Sequence updated (17-Mar-1999). Location/Qualifiers 1..630 /organism="Homo sapiens" /db_xref="taxon:9606" 1..630 /gene="FGF-21" /gene="FGF-21" /codon_start=1 /product="FGF-21" /protein_id="BA99415.1" /db_xref="GI:9049445" /translation="MDSDETEHSLWVSLAGLLGACQAHPIPDSSPLIQFGQV RQRYLYTDAQOQTEAHLEIRDEGTGAGADQSDSESLIQAKLPVQIQIGVYKTSFL CQRPDGLVGSILHDEPEACSFRELLIEDGVNYSFAHGLPLHPNKRPHNDPAPRG PARFLPLGIPALPEPPGIALPPDPVGSSDPLSMVWSQGRSPSYAS" BASE COUNT 117 a 219 c 182 g 112 t ORIGIN

Query Match 61.7%; Score 400.2; DB 85; Length 630; Best Local Similarity 80.6%; Pred. No. 4,1e-89; Matches 466; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 53 tgcgtgctgtctcctgctggtggtgttcaacaagatccccatccctgactcagccccc 112
Db 50 TGCtGGCTGCTCTCTGTGTGGAGCTGCGACGACACCCCATCTGACTCCAGTCTTC 109
QY 113 tctccagtttgggtggtcagtcgcgacagagtacacctctacacagatgacgacaa 172
Db 110 TCCtGCAATTCGGGGGCGCAAGTCGGGAGCGGTCTTACACAGATGATGCCAGCA 169
QY 173 ctgaagccacctgagatcagagagatggaacgtgtgagcgacgacacccgacgtc 232
Db 170 CAGAAGCCCACTTGAGATCAGGAGATGGAGCGTGGGGGCGCTGCTGACCAAGACC 229
QY 233 cagaagctctctgagactcaaaagccttgaagccaggggttcaatcaatcctggtgt 292
Db 230 CCGAAAGTCTCTCGAGCTGAAGACCTTGAAGCCGGGAGTATTCATCTTGGAGTCA 289
QY 293 aagccttaggttctcttgcacaacagacagatgagctctctatgatacgctactttg 352
Db 290 AGACATTCAGGTTCTCTGTCAGCGGCGCCAGATGGGGCCCTGTATGATCGCTCCACTT 349
QY 353 atcctgagcgtcagctcagagacgtcgtcgtgagagacgtttacaatgtgtaccagt 412
Db 350 ACCCTGAGGCTTCAGACTTCGGGAGCTGCTTCTTGAAGAGGATACAAATGTTTACCACT 409
QY 413 ctgaagccatgagcctgcccctgctgtcctcagaaagagactccccaaacagatgcaa 472
Db 410 CCGAAGCCCAAGCGCTCCCGCTGACCTGCCAAGGAGAACAAAGTCCCAACCGGAGACCTG 469
QY 473 catctgtggagcctgtgtgcttctctgcccacagcagcgtctcacaagagcccaagcc 532
Db 470 CACCCGAGACAGCTCTTCTGCTGCACTACAGGAGGCTGCCCCCGCACTCCCGAGGC 529

QY 533 aagcagatctcctgccccagagaccccaagatgtggtcctctctgacccctgacatg 592
Db 530 CACCCGAGATCCTGCGCCCCCAGCCCGAGAGTGGGCTCTCGGACCTCTGAGCATGG 589
QY 593 tagagccttacagggcgaagcccgatgctgtctga 633
Db 590 TGGACCTTCCAGGAGCGAGCCCAAGCTACGCTTCTGTA 630

RESULT 3
AC073828
LOCUS AC073828 215734 bp DNA 29-JUN-2000
DEFINITION Mus musculus clone RP23-9J18, WORKING DRAFT SEQUENCE, 21 unordered pieces.
ACCESSION AC073828
VERSION AC073828.1 GI:8810445
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus.
1 (bases 1 to 215734)
DOE Joint Genome Institute.
Sequencing of Mouse
Unpublished
2 (bases 1 to 215734)
DOE Joint Genome Institute.
Direct Submission
Submitted (29-JUN-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 1740562
Center clone name: RPCI-23_9J18

Summary Statistics
Consensus quality: 19667 bases at least Q40
Consensus quality: 207971 bases at least Q20
Consensus quality: 209928 bases at least Q20
Estimated insert size: 207740; agarose-fp estimation
Estimated insert size: 213734; sum-of-contigs estimation
Quality coverage: 6.12 in Q20 bases; agarose-fp estimation
Quality coverage: 5.94 in Q20 bases; sum-of-contigs estimation.
NOTE: This is a 'working draft' sequence. It currently consists of 21 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
1 1013: contig of 1013 bp in length
* 1014 1113: gap of unknown length
* 1114 2175: contig of 1062 bp in length
* 2176 2275: gap of unknown length
* 2276 3429: contig of 1154 bp in length
* 3430 3529: gap of unknown length
* 3530 4884: contig of 1355 bp in length
* 4885 4984: gap of unknown length
* 4985 6848: contig of 1864 bp in length
* 6849 6948: gap of unknown length
* 6949 9317: contig of 2369 bp in length
* 9318 9417: gap of unknown length
* 9418 12128: contig of 2711 bp in length
* 12129 14970: gap of unknown length
* 14971 15070: contig of 2742 bp in length
* 15071 19571: contig of 4501 bp in length

* 19572 19671: gap of unknown length
* 19672 25961: contig of 6290 bp in length
* 25962 26061: gap of unknown length
* 26062 36753: contig of 10692 bp in length
* 36754 36853: gap of unknown length
* 36854 49116: contig of 12263 bp in length
* 49117 49216: gap of unknown length
* 49217 61879: contig of 12663 bp in length
* 61880 61979: gap of unknown length
* 61980 72513: contig of 10534 bp in length
* 72514 72614: gap of unknown length
* 72614 83821: contig of 11108 bp in length
* 83822 83821: gap of unknown length
* 83822 98633: contig of 14812 bp in length
* 98634 98733: gap of unknown length
* 98734 118323: contig of 19590 bp in length
* 118324 118423: gap of unknown length
* 118424 137668: contig of 19245 bp in length
* 137669 162338: contig of 24570 bp in length
* 162339 162438: gap of unknown length
* 162439 186625: contig of 24187 bp in length
* 186626 186725: gap of unknown length
* 186726 215734: contig of 29009 bp in length.
Location/Qualifiers
1. 215734
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-9J18"
/clone_11b="RPC1 mouse BAC library 23"
BASE COUNT 54325 a 53055 c 52767 g 53583 t 2004 others
ORIGIN

Query Match 45.8%; Score 297.4; DB 75; Length 215734;
Best Local Similarity 99.7%; Pred. No. 3.9e-64;
Matches 298; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 342 gccctacattgatcctcgtgagcctcagctcaggaactcgtcgtgaggaagcgtttacaa 401
Db 136527 GCCTCATTGATCTGAGGCGCTGAGCTTCAAGAACTGCTGAGAGGAGCGGTACAA 136586
QY 402 tctgtaccagctcgtgaagccatcgtgcccctcgtcgtcgtcgtcaggaaggaatcccaaa 461
Db 136587 TGTGACCAAGCTGGAAGCCATGGCGCTGCCCCCTGCTGCTTGCCTCAAGAGACTCCCAAA 136646
QY 462 ccaggaatgcaaatcctcgtggaactcgtgacctcctcctccatgccaagcctgctccaca 521
Db 136647 CCAGGATGCAACATCTGGGGACCTGTGCGCTTCTGCCATGCCAGGCTGCTCCACGA 136706
QY 522 gccccaagaacaaagattcctcgtccccaagagccccaagatgttggtccctcctgaacc 581
Db 136707 GCCCCAAAGCAAGAGATTCTGCCCCCAGAGCCCCCAAGATGTGGGCTCTCTGACCC 136766
QY 582 cctgagcattgtagagcctttacagggccgaagcccaagctatgcgtcctgaactcttc 640
Db 136767 CCTGAGCATGTGTAGAGCTTTACAGGGCCCAAGCCCAAGCTATGCTGACTCTTCC 136825

RESULT 4
AF214655/c 9257 bp DNA ROD 23-MAR-2000
LOCUS Mus musculus alpha(1.2) fucosyltransferase FUT1 (Fut1) gene.
DEFINITION complete cds.
ACCESSION AF214655
VERSION AF214655.1 GI:7288502
KEYWORDS house mouse.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 9257)
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Domino, S.E. and Lowe, J.B.

TITLE Mus musculus alpha(1.2)fucosyltransferase FUT1 genomic sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 9257)
AUTHORS Domino, S.E. and Lowe, J.B.
TITLE Direct Submission
JOURNAL Submitted (13-DEC-1999) HHMI, University of Michigan, 1150 W. Med.
Ctr. Dr., Ann Arbor, MI 48109-0650, USA
FEATURES
Location/Qualifiers
1. 9257
/organism="Mus musculus"
/strain="129/Ola"
/db_xref="taxon:10090"
/chromosome="7"
/map="23.2: between K1k1 and Ldh3"
/cell_line="E14"
/cell_type="ES"
/gene="Fut1"
/product="alpha(1.2)fucosyltransferase FUT1"
/EC_number="2.4.1.69"
/note="glycosyltransferase: similar to the Mus musculus strain ICR sequences deposited at GenBank Accession Numbers Y09683 and AF113533; similar to the Mus musculus strain NIH Swiss sequence deposited at GenBank Accession Number U90553; similar to the Homo sapiens H blood group alpha(1.2)fucosyltransferase FUT1 sequences deposited at GenBank Accession Numbers M3531 and Z69587"
/codon_start=1
/product="alpha(1.2)fucosyltransferase FUT1"
/protein_id="AAE45145.1"
/db_xref="GI:7288503"
/translation="MMPSSRRQLCTFLVLCVLSGSPFFHLNGNFFRNGLTSLVLC SYHLKSPVAMVCLPPILOTNSGSPCEGSSLSGTTTPGRRGNOMGOVATLL ALAOLNGROAFIOPMHAALAPVPRISIPVDIPVDSITPMOHLVLDMSSEYSHLE DPFLLSGPFCSSWTFEHLRLREDFLHLHRLDGAQYLLSGIRIPAGIRPTFPG VHVRRGDYLEVPMNRKGVGDRAVLQDAMFRRHRDPLFVYTSNGMKMCLENTDT SHGDVFEAGNGDEGTPGDFALITQCNFTIMTIGTFGMAVYLAGGDTVYLANTLRPD SEFLKIFRPEAAFLTEWVGINADLSPILOAQDPWKPDSLFRV"

polyA_signal
BASE COUNT 2267 a 2195 c 2580 g 2215 t
ORIGIN
Query Match 45.7%; Score 296.6; DB 94; Length 9257;
Best Local Similarity 98.7%; Pred. No. 1.2e-63;
Matches 299; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 342 gccctacattgatcctcgtgagcctcagctcaggaactcgtcgtgaggaagcgtttacaa 401
Db 1354 GCCTCATTGATCTGAGGCTGAGCTTCAAGAACTGCTGAGAGGAGGTTACAA 1295
QY 402 tctgtaccagctcgtgaagccatcgtgacctcctcctccatgccaagcctgctccaca 461
Db 1294 TGTGACCAAGCTGGAAGCCCAAGGCGCTGCCCCCTGCTGCTCAAGAGGACTCCCAAA 1235
QY 462 ccaggaatgcaaatcctcgtggaactcgtgacctcctcctccatgccaagcctgctccaca 521
Db 1234 CCAGGATGCAACATCTGGGGACCTGTGCGCTTCTGCCATGCCAGGCTGCTCCACGA 1175
QY 522 gccccaagaacaaagattcctcgtccccaagagccccaagatgttggtccctcctgaacc 581
Db 1174 GCCCCAAAGCAAGAGATTCTGCCCCCAGAGCCCCCAAGATGTGGGCTCTCTGACCC 1115
QY 582 cctgagcattgtagagcctttacagggccgaagcccaagctatgcgtcctgaactcttc 641
Db 1114 CCTGAGCATGTGTAGAGCTTTACAGGGCCCAAGCCCAAGCTATGCTGACTCTTCC 1055
QY 642 tga 644
||

	DB	1054	GGA	1052	
RESULT	5				
AX097639					
LOCUS	AX097639	514 bp	DNA	PAT	30-MAR-2001
DEFINITION	Sequence 1 from Patent WO0118209.				
ACCESSION	AX097639				
VERSION	AX097639.1	GI:13514271			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Shinkels,R.A., Vernet,C., Burgess,C., Fernandes,E., Taupier,R.J., Quinn,K.E., Spyteck,K.A., Rastelli,L. and Herrmann,J.L. Fibroblast growth factor polypeptide and nucleic acids encoding same				
TITLE	Patent: WO 0118209-A 1 15-MAR-2001;				
JOURNAL	Curagen Corporation (US)				
FEATURES	Location/Qualifiers				
SOURCE	1..514				
CDS	/organism="Homo sapiens" /db_xref="taxon:9606" 52..513 /note="unnamed protein product" /codon_start=1 /protein_id="CAC35437.1" /db_xref="GI:13514272" /translation="MDSDTEFEHSGIWMVSYLAGLLGACQAMPIDPSSPLLOFGGVN KORLYTDDAQCTEHLREDDTGVGAADSPSLQLKALKEGVIOIGVKTSRFL CQRDGLAYGSILHEDPEASFRLELLDGEDYWNVOSEAHGLPLHLPGHRL"				
BASE COUNT	103 a 157 c 155 g 99 t				
ORIGIN					
Query Match	43.7%; Score 283.8; DB 10; Length 514;				
Best Local Similarity	82.9%; Pred. No. 3.5e-60;				
Matches 324; Conservative	0; Mismatches 67; Indels 0; Gaps 0;				
Oy	53	tgcgtgctgtcctcctgctgggagtcacaaagcatccatccctactccagcccc	112		
Db	101	tgcitgctgctgcttctgcttggaacctcccaaggcacaccattccctactcctcgc	160		
Oy	113	tccctcagtttggygttaagtccgycagaaggttaccttacaagatagcagaagaca	172		
Db	161	tctctcattgtcgggggccaagtcgccggcagcggtacctctacacagatgatgccacgaca	220		
Oy	173	ctgaagcccacatggaagtcagggagagtgtaaagctgtgtaggcagacacacgcagtc	232		
Db	221	cagaagccctacatggagatcagggagtgagcgtggggggcctctctgacacagagcc	280		
Oy	233	cagaagctcctctgagagtcacaagccttgaaagccagggggtcatccaatccttggtgtca	292		
Db	281	ccgaagagctctctgcagctgaaagccttgaaagccggaggttatccaatcttggagagca	340		
Oy	293	aagcctctaagttctcttggccaacagccagatgtagctctctatagatcgctcaattg	352		
Db	341	agacattccagagttctcttggtgcacagccacagatggggccctctatgagatcgctccaccttg	400		
Oy	353	atcctgagccttcagagctcagaagagctgctcgtcagaggaagcgtttacaatgtgtacaggt	412		
Db	401	accctgagccttcagagctcagaagagctgctcgtcagaggaagcgtttacaatgtgtacaggt	460		
Oy	413	ctgaagcccatgctgcctccctgcgctgc 443			
Db	461	ccgaagcccatgctgcctccctgcgctgc 491			
RESULT	6				
AC024740	AC024740	179538 bp	DNA	HTG	07-JUL-2000
LOCUS					

DEFINITION	Homo sapiens chromosome 19 clone RP11-801D6, WORKING DRAFT
SEQUENCE	17 unordered pieces.
ACCESSION	AC024740
VERSION	AC024740.3
KEYWORDS	GI:8570405
SOURCE	HTG; HTGS_PHASE1; HTGS_DRAFT.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 179538)
AUTHORS	Waterston,R.H.
TITLE	The sequence of Homo sapiens clone
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 179538)
TITLE	Waterston,R.H.
REFERENCE	Submitted (01-MAR-2000) Genome Sequencing Center, Washington
AUTHORS	University School of Medicine, 4444 Forest Park Parkway, St. Louis
TITLE	MO 63108, USA
COMMENT	On Jun 17, 2000 this sequence version replaced gi:7263917.

Center: Washington University Genome Sequencing Center	Genome Center
Center code: WUGSC	
Web site: http://genome.wustl.edu/gsc/index.shtml	
Project Information	
Center project name: H.NH0801D06	
Summary Statistics	
Sequencing vector: M13, 100%	
Sequencing vector: plasmid: 0%	
Chemistry: Dye-primer ET; 97% of reads	
Chemistry: Dye-terminator Big Dye; 3% of reads	
Assembly program: Phrap; version 0.990319	
Consensus quality: 173830 bases at least Q40	
Consensus quality: 175557 bases at least Q30	
Consensus quality: 176446 bases at least Q20	
Insert size: 242000; agarose-fp	
Insert size: 177938; sum-of-contigs	
Quality coverage: 5.32 in Q20 bases; agarose-fp	
Quality coverage: 6.01 in Q20 bases; sum-of-contigs	

* NOTE: This is a 'working draft' sequence. It currently

* consists of 17 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

1	1801: contig of 1801 bp in length
*	
*	1802
*	1901: gap of unknown length
*	1902
*	4641: contig of 2740 bp in length
*	4642
*	4741: gap of unknown length
*	4742
*	9679: contig of 4938 bp in length
*	9680
*	9779: gap of unknown length
*	9780
*	13392: contig of 3613 bp in length
*	13393
*	13492: gap of unknown length
*	13493
*	17164: contig of 3672 bp in length
*	17165
*	17264: gap of unknown length
*	17265
*	21510: contig of 4246 bp in length
*	21511
*	21610: gap of unknown length
*	21611
*	27089: contig of 5479 bp in length
*	27090
*	27189: gap of unknown length
*	27190
*	33636: contig of 8447 bp in length
*	35637
*	35736: gap of unknown length
*	45537: contig of 9801 bp in length
*	45538
*	45637: gap of unknown length
*	45638
*	45651: contig of 10814 bp in length
*	45652
*	56551: gap of unknown length
*	56552
*	66802: contig of 10251 bp in length
*	66803
*	66902: gap of unknown length
*	79118: contig of 12216 bp in length
*	79119
*	79218: gap of unknown length

QY	DB	Query Match	Best Local Similarity	Matches	256; Conservative	0;	Mismatches	98;	Indels	0;	Gaps	0;
QY	280	atccgaggtgtcaaaagcctctagttctttcttgccaacagccagatgagacctctctatgga	30.4%;	Score 197.2;	DB 68;	Length 179538;						
DB	130126	ATCCGTGGGCTTTACATCAGGAAACAGAGAACCCGTCTGTGATCCGTGTTTGTGCCCC				130185						
QY	340	tgcgctcaattgataccctcctcagcctcgcagcctccagagaactctgctgagagcgttac				399						
DB	130186	TAGCTCCACATTTGACCCCTGAGGCTCTGACACTTCCGGGAGCTGCTTCTTAGGACAGATATC				130245						
QY	400	aatgtgtaccagctctgtaagcccatcgtgctgcctctgctgctctcctcagaagactcccca				459						
DB	130246	AATGTATTACATGTCGGAAGCCACAGGCGCTCCGGCTGACACTGTGCAAGGGAACAAGTCCCA				130305						
QY	460	aaccagatgcacacatcctcgtggagcctgtggtcttcctgcgcctatgccagcctgtctccac				519						

```

Db 130306 CAGCGGGAGACCCCTGCACACCCGAGGACACAGTGCCTTCCTGCGACACACGAGGCTCCGCCCC 130365
Oy 520 gagcccccaagaccagcagatctctgcccccaagagcccccaagatgtagctctctgac 579
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 130366 GCACCCCCGGAGCACCCGGAAATCCTTGCGCCCCCAGCCGCCCCCGAGATGCGGCTCTCGGAC 130425
Oy 580 cccctgagcatgtagagcagccttaacggcgcaagcccccaagatgtagctctga 633
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 130426 CCTGTGAGCATGtGGAGACCTTCTCCAGGGCCGAGACCCCGACGTATGCTTCTGA 130479

RESULT 7
AC008749.0/c
MPCOMMENT
Sequence split into 5 fragments LOCUS AC008749 Accession AC008749
Fragment Name Begin End
AC008749..0 1 110000
AC008749..1 100001 210000
AC008749..2 200001 310000
AC008749..3 300001 410000
AC008749..4 400001 467420
LOCUS AC008749 467420 bp DNA HTG 21-JUN-2000
DEFINITION Homo sapiens chromosome 19 clone CTD-2611012, WORKING DRAFT
ACCESSION AC008749
VERSION AC008749.5 GI:8575912
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 467420)
            DOE Joint Genome Institute.
            Sequencing of Human Chromosome 19
            Unpublished
            2 (bases 1 to 467420)
            DOE Joint Genome Institute.
            Direct Submission
            Submitted (03-NOV-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
            On Jun 21, 2000 this sequence version replaced gi:7689801.
            -----Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -----
            Project Information
            Center Project Name: 832953, BC797977
            Center clone name: CITB-EL_2611012
            -----
            Summary Statistics
            Consensus quality: 353446 bases at least Q40
            Consensus quality: 412781 bases at least Q30
            Consensus quality: 426617 bases at least Q20
            Estimated insert size: 151170; agarose-fp estimation
            Estimated insert size: 454620; sum-of-contigs estimation
            Quality coverage: 9.78 in Q20 bases; agarose-fp estimation
            Quality coverage: 3.25 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 127 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            1 1208: contig of 1208 bp in length
            * 1209 1308: gap of unknown length
            * 1309 2441: contig of 1133 bp in length
            * 2442 2541: gap of unknown length
            * 2542 3647: contig of 1106 bp in length
            * 3648 3747: gap of unknown length
            * 3748 4788: contig of 1041 bp in length

```


* 4789 4888: gap of unknown length
* 4889 6132: contig of 1244 bp in length
* 6133 6232: gap of unknown length
* 6233 7544: contig of 1312 bp in length
* 7544 7644: gap of unknown length
* 7645 8892: contig of 1248 bp in length
* 8893 8992: gap of unknown length
* 8992 10296: contig of 1304 bp in length
* 10297 10397: gap of unknown length
* 10397 11725: contig of 1328 bp in length
* 11726 11825: gap of unknown length
* 11826 13734: contig of 1909 bp in length
* 13735 13834: gap of unknown length
* 13835 15056: contig of 1222 bp in length
* 15057 15156: gap of unknown length
* 15157 16358: contig of 1202 bp in length
* 16359 16458: gap of unknown length
* 16459 17875: contig of 1417 bp in length
* 17876 17975: gap of unknown length
* 17976 19527: contig of 1552 bp in length
* 19528 19627: gap of unknown length
* 19628 21424: contig of 1797 bp in length
* 21425 21524: gap of unknown length
* 21525 22793: contig of 1269 bp in length
* 22794 22893: gap of unknown length
* 22894 23905: contig of 1012 bp in length
* 23906 24005: gap of unknown length
* 24006 25440: contig of 1435 bp in length
* 25441 25540: gap of unknown length
* 25541 26895: contig of 1355 bp in length
* 26896 26995: gap of unknown length
* 26996 28178: contig of 1183 bp in length
* 28179 28278: gap of unknown length
* 28279 29887: contig of 1609 bp in length
* 29888 29987: gap of unknown length
* 29988 31869: contig of 1882 bp in length
* 31870 31969: gap of unknown length
* 31970 33535: contig of 1566 bp in length
* 33536 33635: gap of unknown length
* 33636 34963: contig of 1328 bp in length
* 34964 35063: gap of unknown length
* 35064 36881: contig of 1818 bp in length
* 36882 36981: gap of unknown length
* 36982 38148: contig of 1167 bp in length
* 38149 38248: gap of unknown length
* 38249 39889: contig of 1641 bp in length
* 39890 39989: gap of unknown length
* 39990 41963: contig of 1974 bp in length
* 41964 42063: gap of unknown length
* 42064 43733: contig of 1670 bp in length
* 43734 43833: gap of unknown length
* 43834 45414: contig of 1581 bp in length
* 45415 45514: gap of unknown length
* 45515 47046: contig of 1532 bp in length
* 47047 47146: gap of unknown length
* 47147 48965: contig of 1819 bp in length
* 48966 49065: gap of unknown length
* 49066 51369: contig of 2304 bp in length
* 51370 51469: gap of unknown length
* 51470 53128: contig of 1659 bp in length
* 53129 53228: gap of unknown length
* 53229 54480: contig of 1252 bp in length
* 54481 54580: gap of unknown length
* 54581 56485: contig of 1905 bp in length
* 56486 56585: gap of unknown length
* 56586 58664: contig of 2079 bp in length
* 58665 58764: gap of unknown length
* 58765 59959: contig of 1195 bp in length
* 59960 60059: gap of unknown length
* 60060 62303: contig of 2244 bp in length
* 62304 62403: gap of unknown length
* 62404 64204: contig of 1801 bp in length
* 64205 64304: gap of unknown length

* 64305 66196: contig of 1892 bp in length
* 66197 66296: gap of unknown length
* 66297 67878: contig of 1582 bp in length
* 67879 67978: gap of unknown length
* 67979 69965: contig of 1987 bp in length
* 69966 70065: gap of unknown length
* 70066 71633: contig of 1568 bp in length
* 71634 71734: gap of unknown length
* 71734 73978: contig of 2245 bp in length
* 73979 74078: gap of unknown length
* 74079 76073: contig of 1995 bp in length
* 76074 76173: gap of unknown length
* 76174 78469: contig of 2296 bp in length
* 78470 78569: gap of unknown length
* 78570 81250: contig of 2681 bp in length
* 81251 81350: gap of unknown length
* 81351 83762: contig of 2412 bp in length
* 83763 83862: gap of unknown length
* 83863 85889: contig of 2127 bp in length
* 85890 86089: gap of unknown length
* 86090 88156: contig of 2067 bp in length
* 88157 88256: gap of unknown length
* 88257 90318: contig of 2062 bp in length
* 90319 90418: gap of unknown length
* 90419 92289: contig of 1871 bp in length
* 92290 92389: gap of unknown length
* 92390 95174: contig of 2785 bp in length
* 95175 95274: gap of unknown length
* 95275 98333: contig of 3059 bp in length
* 98334 98433: gap of unknown length
* 98434 99975: contig of 1542 bp in length
* 99976 100075: gap of unknown length
* 100076 101499: contig of 1424 bp in length
* 101500 101599: gap of unknown length
* 101600 103848: contig of 2249 bp in length
* 103849 103948: gap of unknown length
* 103949 106127: contig of 2179 bp in length
* 106128 106227: gap of unknown length
* 106228 108402: contig of 2175 bp in length
* 108403 108502: gap of unknown length
* 108503 110948: contig of 2446 bp in length
* 110949 111048: gap of unknown length
* 111049 113905: contig of 2857 bp in length
* 113906 114005: gap of unknown length
* 114006 117257: contig of 3251 bp in length
* 117257 117356: gap of unknown length
* 117357 119135: contig of 1779 bp in length
* 119136 119235: gap of unknown length
* 119236 121325: contig of 2090 bp in length
* 121326 121425: gap of unknown length
* 121426 123954: contig of 2529 bp in length
* 123955 124054: gap of unknown length
* 124055 126851: contig of 2797 bp in length
* 126852 126951: gap of unknown length
* 126952 129652: contig of 2701 bp in length
* 129653 129752: gap of unknown length
* 129753 132233: contig of 2481 bp in length
* 132234 132333: gap of unknown length
* 132334 135268: contig of 2935 bp in length
* 135269 135368: gap of unknown length
* 135369 137746: contig of 2378 bp in length
* 137747 137846: gap of unknown length
* 137847 141549: contig of 3703 bp in length
* 141550 141649: gap of unknown length
* 141650 144391: contig of 2742 bp in length
* 144392 144491: gap of unknown length
* 144492 147414: contig of 2923 bp in length

Query Match 30.1%; Score 195.6; DB 60; Length 110000;
Best Local Similarity 72.0%; Pred. No. 8; 8e-39;
Matches 255; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 280 atccctggttcaaaagcctctagttcttcttgcacaaagccagatgagctctctatgga 339

```

|||||
Db 109817 ATCTGGGCTTACATCAGGAAACAGAGAACCTGTCTCTGATCCTGTTTGTCCCC 109758
QY 340 tcgctcaatttgatcctgagcctgctcagatctcagagaactgctgtagagaggttac 399
Db 109757 TAGCTCCACTTTGACCTCGAGGCGCTTGCGAGCTGCTCTTGAGAGCGGATAC 109698
QY 400 aatgtgtacagatctgagagccatgagcctgcccctgctgctgctcagagaagactccca 459
Db 109697 AATGTTTACAGTCCGAGCCACGCGCTCCCGCTGACCTGACAGGAGAACAGTCCCA 109638
QY 460 aaccagagatgaacatctctgtagctgtgcttctcctcctcagagccagcctctccac 519
Db 109637 CACCGGGACCTGACACCCGAGGACCACTGCTTCTGCTGACACAGGCGCTCCGCC 109578
QY 530 gagccccaagacccaagagagatctcctgccccagagccccaagatgtgagctcctgac 579
Db 109577 GCACCCCGGAGCCACCGGAGATCTGCGCCCGACGCCCCCGATGTGGCTCTCGGAC 109518
QY 580 cccctgagcatgtgtagagcctttacagagccgaagccccaagctatgctgctga 633
Db 109517 CCTCTGAGCATGTGGAGCACTTCCACGAGGCGGAGCCGACCTACGCTTCTCTGA 109464

```

RESULT 8 AC008749_1/c WPCOMMENT

Sequence split into 5 fragments LOCUS AC008749 Accession AC008749

Fragment Name	Begin	End
AC008749_0	1	110000
AC008749_1	100001	210000
AC008749_2	200001	310000
AC008749_3	300001	410000
AC008749_4	400001	467420

Continuation (2 of 5) of AC008749 from base 100001 (AC008749 Homo sapiens chromosome 19)

Query Match 30.1%; Score 195.6; DB 60; Length 110000;
Best Local Similarity 72.0%; Pred. No. 8.8e-39;
Matches 255; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

```

QY 280 atctcgggtgtcaagacccctagttctcttgcacaagcagaatgagctctatgga 339
Db 9817 ATCTGGGCTTACATCAGGAAACAGAGAACCTGTCTCTGATCCTGTTTGTCCCC 9758
QY 340 tcgctcaatttgatcctgagcctgctcagatctcagagaactgctgtagagaggttac 399
Db 9757 TAGCTCCACTTTGACCTCGAGGCGCTTGCGAGCTGCTCTTGAGAGCGGATAC 9698
QY 400 aatgtgtacagatctgagagccatgagcctgcccctgctgctgctcagagaagactccca 459
Db 9697 AATGTTTACAGTCCGAGCCACGCGCTCCCGCTGACCTGACAGGAGAACAGTCCCA 9638
QY 460 aaccagagatgaacatctctgtagctgtgcttctcctcctcagagccagcctctccac 519
Db 9637 CACCGGGACCTGACACCCGAGGACCACTGCTTCTGCTGACACAGGCGCTCCGCC 9578
QY 520 gagccccaagacccaagagagatctcctgccccagagccccaagatgtgagctcctgac 579
Db 9577 GCACCCCGGAGCCACCGGAGATCTGCGCCCGACGCCCCCGATGTGGCTCTCGGAC 9518
QY 580 cccctgagcatgtgtagagcctttacagagccgaagccccaagctatgctgctga 633
Db 9517 CCTCTGAGCATGTGGAGCACTTCCACGAGGCGGAGCCGACCTACGCTTCTCTGA 9464

```

RESULT 9
AB006136/c
LOCUS AB006136 4033 bp DNA PRI 20-JUL-2000
DEFINITION Homo sapiens gene for alpha 1,2-fucosyltransferase, 5' flanking
region and partial cds.
ACCESSION AB006136
VERSION AB006136.1 GI:3242453

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
TITLES
JOURNAL
COMMENT
FEATURES
source

FUT1; alpha (1, 2) fucosyltransferase.
 Homo sapiens peripheral leukocytes DNA.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (sites)
 Koda, Y., Soejima, M. and Kimura, H.
 Changing transcription start sites in H-type
 alpha(1,2)fucosyltransferase gene (FUT1) during differentiation of
 the human erythroid lineage
 Eur. J. Biochem. 256 (2), 379-387 (1998)
 98430978
 2 (bases 1 to 4033)
 Koda, Y.
 Direct Submission
 Submitted (04-AUG-1997) to the DDBJ/EMBL/GenBank databases. Yoshio
 Koda, Kurume University, School of Medicine, Department of Forensic
 Medicine; Asahimachi 67, Kurume, Fukuoka 830-0011, Japan
 (E-mail: ykoda@med.kurume-u.ac.jp, Tel: 81-942-31-7554,
 Fax: 81-942-31-7700)
 On Jun 20, 1998 this sequence version replaced gi:2317262.
 Sequence updated (16-Jun-1998).
 Location/Qualifiers
 1..4033
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="19"
 /map="19q13.3"
 /rname="FUT1"
 /rsize_type="peripheral leukocytes"
 1..510
 511..578
 /gene="FUT1"
 /number=1
 511..578
 /gene="FUT1"
 511..578
 /gene="FUT1"
 /standard_name="alpha(1,2)fucosyltransferase"
 /codon_start=1
 /product="FUT1"
 /protein_id="BA28952.1"
 /db_xref="GI:3242454"
 /translation="MGFHHVQAGLEILTSGLPAWY"
 BASE COUNT 877 a 1242 c 1099 g 815 t
 ORIGIN

Query Match 29.9%; Score 194; DB 85; Length 4033;
Best Local Similarity 71.8%; Pred. No. 4.5e-38;
Matches 254; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

```

QY 280 atctcgggtgtcaagacccctagttctcttgcacaagcagaatgagcctctatgga 339
Db 1593 ATCTGGGCTTACATCAGGAAACAGAGAACCTGTCTCTGATCCTGTTTGTCCCC 1534
QY 340 tcgctcaatttgatcctgagcctgctcagatctcagagaactgctgtagagaggttac 399
Db 1533 TAGCTCCACTTTGACCTCGAGGCGCTTGCGAGCTGCTCTTGAGAGCGGATAC 1474
QY 400 aatgtgtacagatctgagagccatgagcctgcccctgctgctcagagaagactccca 459
Db 1473 AATGTTTACAGTCCGAGCCACGCGCTCCCGCTGACCTGACAGGAGAACAGTCCCA 1414
QY 460 aaccagagatgaacatctctgtagctgtgcttctcctcctcagagccagcctctccac 519
Db 1413 CACCGGGACCTGACACCCGAGGACCACTGCTTCTGCTGACACAGGCGCTCCGCC 1354
QY 520 gagccccaagacccaagagagatctcctgccccagagccccaagatgtgagctcctgac 579
Db 1353 GCACCCCGGAGCCACCGGAGATCTGCGCCCGACGCCCCCGATGTGGCTCTCGGAC 1294
QY 580 cccctgagcatgtgtagagcctttacagagccgaagccccaagctatgctgctga 633
Db 1294 cccctgagcatgtgtagagcctttacagagccgaagccccaagctatgctgctga 633

```

Db 1293 CCTGTGACATGTGGACCTTCCAGGCGCGAAGCCCGACTACGCTTCTTA 1240

RESULT 10
AC009002/c 37402 bp DNA PRI 13-AUG-2000
LOCUS Homo sapiens chromosome 19 clone L1NLR-279G3, complete sequence.
DEFINITION AC009002
ACCESSION AC009002.5 GI:9799769
VERSION
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 37402)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 37402)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 37402)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (13-AUG-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Aug 13, 2000 this sequence version replaced gi:9256105.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.
STS Content:
SHGC-35310 G28569.
FEATURES
Location/Qualifiers
source 1..37402
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="L1NLR-279G3"
BASE COUNT 8189 a 9990 c 10374 g 8849 t
ORIGIN

Query Match 29.9%; Score 194; DB 87; Length 37402;
Best Local Similarity 71.8%; Pred. No. 2.8e-38;
Matches 254; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 280 atccctgggtgtcaagcctctagttcttctgccaagcagatgagctctctatga 339
|||||
Db 10061 ATCCTGGGTTCACATCAGAAAACAGAGAACCTGTCTCTATCTCTTTTGTGCC 10002
|||||

QY 340 tgcgctcaacttctcctgagcctgagctcagctcagaagactgctgtgagagcgttac 399
|||||
Db 10001 TACGTCACATTGACCTGAGGCTGACGCTCCGAGAGCTGCTTGTGAGCGGATAC 9942
|||||

QY 400 aatggttaccagctgaaagccatgctgcctcctcgtctgctcctaagaagaactccca 459
|||||
Db 9941 AATGTTTACAGTCCGAAGCCCAAGGCTCCCTCCCTACCTGCGAAGCAAGTCCCA 9882
|||||

QY 460 aaccagatcacatcctctggaactgtgcctctctccatgccaagcctgctccac 519
|||||
Db 9881 CACCGGACCTTCACACCCGAGACCAAGCTTCCTCTGCTACTACCAAGCCTGCCCCC 9822
|||||

QY 520 gagccccaagaacacagagatctctgcctcccaagcctcccaagatgtggctcctcga 579
|||||
Db 9821 GCACCTCCGAGAGCCACCGCAATCTGCGCCCGCCCGCCGATGTGGGCTCTCGGAC 9762
|||||

QY 580 cccctagacatgttagagccttaagggcggaagcccaagctatgctcctga 633
|||||

Db 9761 CCTGTGACATGTGGACCTTCCAGGCGCGAAGCCCGACTACGCTTCTTA 9708

RESULT 11
AC026803/c 227949 bp DNA HTG 18-JUL-2000
LOCUS Homo sapiens chromosome 19 clone CTD-2639E6, WORKING DRAFT
DEFINITION AC026803
ACCESSION AC026803
VERSION AC026803.4 GI:9256691
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 227949)
AUTHORS Sequencing of Human Chromosome 19
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 227949)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (24-MAR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jul 18, 2000 this sequence version replaced gi:8576168.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 843459, BC808483
Center clone name: CTRB-EL_2639E6

Summary Statistics
Consensus quality: 217198 bases at least Q40
Consensus quality: 224116 bases at least Q30
Consensus quality: 225875 bases at least Q20
Estimated insert size: 226670; agarose-fp estimation
Estimated insert size: 227399; sum-of-contigs estimation
Quality coverage: 6.85 in Q20 bases; agarose-fp estimation
Quality coverage: 6.83 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

1 21840: contig of 21840 bp in length
1 21841: gap of unknown length
* 21941: 31559: contig of 9619 bp in length
* 31560: gap of unknown length
* 31687: contig of 4528 bp in length
* 36288: gap of unknown length
* 59202: contig of 22915 bp in length
* 59203: gap of unknown length
* 59303: 69077: contig of 9775 bp in length
* 69078: gap of unknown length
* 69178: 80927: contig of 11750 bp in length
* 80928: gap of unknown length
* 81028: 138192: contig of 57165 bp in length
* 138193: gap of unknown length
* 138293: 160224: contig of 21932 bp in length
* 160225: gap of unknown length
* 160325: 161182: contig of 858 bp in length
* 161183: gap of unknown length
* 161283: 162861: contig of 1579 bp in length
* 162862: gap of unknown length
* 162962: 221619: contig of 58658 bp in length
* 221620: gap of unknown length

FEATURES	source
*	221120 225761: contig of 4042 bp in Length
*	225762 225861: gap of unknown length
*	225862 227949: contig of 2088 bp in Length
	Location/Qualifiers
	1. 227949

	/clone_11b="CaITech human BAC library D"									
BASE COUNT	54671	a	58658	c	58905	g	54511	t	1204	others
ORIGIN										

Query Match	20.88;	Score 135.2;	DB 70;	Length 227949;
Best Local Similarity	82.48;	Pred. No. 7.5e-24;		
Matches 155; Conservative	0;	Mismatches 33;	Indels 0;	Gaps 0;

QY	53	tgagagctgctctctcgtcgtgaggtctaccagaatcccatcccgatccagcccc	112
Db 227632	TGCTGGCTGGCTCTCTGTGTGGAGCTTGGCCAGGCACACCCATCCTGACTCCAGTCTC	227573	
QY	113	tcctccagtttgggggtcgaagtcctcggcagaagtacacctcacagatgacagccaagaca	172
Db 227572	TCCTGCATTTGGGGGGCCAGATCCGGGAGGGGTACCTCTACACAGATGATGCCACAGAGA	227513	
QY	173	ctgaagccaccctgagatcaggaagatggaacagtgtgtagcgcaagcaaccgcaatc	232
Db 227512	CAGAAAGCCCACTGGAGATCAGGAGGATGGAGGTGGGGGGCGCTGTCACACAGAGCC	227453	
QY	233	cagaaggt 240	
Db 227452	CCGAAAGT 227445		

RESULT	12
AB037973	
LOCUS	AB037973
DEFINITION	Homo sapiens FGF-23 mRNA for FGF-23,
ACCESSION	AB037973
VERSION	AB037973.1 GI:10118773
KEYWORDS	.
SOURCE	Homo sapiens cDNA from mRNA.
ORGANISM	Homo sapiens

REFERENCE	AUTHORS	TITLE
1 (sites)	Yoshioka, M. and Itoh, N.	Identification of a novel fibroblast growth factor, FGF-23, preferentially expressed in the ventrolateral thalamic nucleus of the brain

```

      /db_xref="GI:10119774"
      /translation="MLGARLRLWYCALCVSCNSVLRAYPNASPLLGSSMGLLHLVY
      ATARSTYLQIHKNGHVDGAPHOTIYSALIRINSEDGFAVITIGVSRNPLICMDGRGSLN
      FGSHTFDPENCRFQHOTLENGDYDYSHPCHFLVSGRKARLWYKNNRPEYSCDRLT
      REIDILPHNTPIRRHRTSRSGEDDSEDPPLNTVAPASCSDELPAEDNENST
      MASDELGVARRGRNTHAGTGSGDCCGPEPFKFI"
      BASE COUNT      159      a      250      c      212      g      135      t

```

Query Match	10.5%;	Score 68.4;	DB 85;	Length 756;
Best Local Similarity	53.8%;	Pred. No. 1e-06;		
Matches 141;	Conservative	0;	Mismatches 211;	Indels 0;
				Gaps 0;

Qy 180 ccacctgagatcagggaagtgtgcagcgacaccgcgagtcagaag 239
||||| |||| | |||| |||| | |||
Db 153 ccacctgagattccacaagaatggccatgtgatggccgccaccctatcacaccaatcttacag 212

OY	240	tccctcgaagctcaaaacgcttggagccaaggggtatctcaatccctgggttcacaagctc	259
Db	213	TGccCTATATGATTCAGATCAGAGGATGTGGCTTTGTGTGATTACAGCTGTGAATGACAG	272
OY	300	taggtttcttcgccacaacaccagaatggagtctctatgtagcgctcacttgatcctga	358
Db	273	AAGATTAACCTTGCAATGGATTTCAGAGGCAACAATTTTGGATCACACTATTTCCGCCGGA	332
OY	360	ggacctgaagcttcagagaactgcgtcgtctggagaagcgggttaacatgtygtaccagttctgaagc	419
Db	333	GAACTGCAGGGTTCCAAACCACGAGCGCTGGAAAACGGGTAGACAGCTTACCACTTCCTCCA	392
OY	420	ccatggcgtgcacctggctgctc	441
Db	393	GTATACACTTCTCGTCACTGTG	414

[illegible]

REFERENCE
AUTHORS
TITLE

1 (bases 1 to 1612)
White, K.E., Evans, W.E., O'Riordan, J.L.H., Speer, M.C., Econs, M.J.,
Lorenz-Deplaigne, B., Grabowski, M., Melinger, T. and Strom, T.M.
Autosomal dominant hypophosphataemic rickets is associated with
mutations in *FGF23*

```

gene      1..1612
          /gene="FGF23"
CDS       147..902
          /gene="FGF23"
          /codon_start=1
          /product="FGF23"
          /protein_id="AA009917.1"

```

```
protein_id="AAG09917.1"
```


